University of Alberta

Structural, Functional, and Thermodynamic Studies of Protein Phosphatase-1 Inhibitors

By

John Randy Bagu ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment
of the requirements for the degree of Doctor of Philosophy

Department of Biochemistry

Edmonton, Alberta

Fall 1999

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-46803-8

Canada

University of Alberta

Library Release Form

Name of Author: John Randy Bagu

Title of Thesis: Structural, Functional, and Thermodynamic Studies of Protein Phosphatase-1 Inhibitors

Degree: Doctor of Philosophy

Year this Degree Granted: 1999

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

John Bagu

John Bagu
11716 8th Avenue
Edmonton, Alberta
T6J 6Z8

Date: October 4, 1999

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommended to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled Structural, Functional, and Thermodynamic Studies of Protein Phosphatase-1 Inhibitors in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Brian Sykes
Dr. Brian Sykes

Charles B Holmes
Dr. Charles Holmes

Dr. Bob Hodges

George Kotovych
Dr. George Kotovych

E Meiering
Dr. Elizabeth Meiering

Date: July 30, 1999

Abstract

Serine/threonine protein phosphatases are responsible for maintaining one half of the regulation of reversible protein phosphorylation which is a control mechanism for many cellular processes. These phosphatases can be subdivided into different groups including protein phosphatase-1 and protein phosphatase-2B (calcineurin). Normal regulation in cells of protein phosphatase-1 is by the protein inhibitor-1. Marine toxins such as microcystins, nodularins (motuporin), okadaic acid, and calyculins are also able to inhibit protein phosphatase-1 leading to liver toxicity and tumor promotion.

The free solution structures of microcystin-LR and motuporin show structural similarities that account for their functional similarity in inhibiting protein phosphatase-1. A difference in their structures explains why microcystin-LR is able to covalently link to protein phosphatase-1, while motuporin lacks this ability. Microcystin-LL, another member of the microcystin family, resembles microcystin-LR. Okadaic acid and calyculin A despite having diverse primary structures are like microcystin and motuporin in tertiary structure accounting for their ability to inhibit protein phosphatase-1. The bound form of microcystin-LR is almost identical to the free form indicating that this inhibitor does not undergo major conformational changes upon binding.

Inhibitor-1 requires phosphorylation to become an active inhibitor of protein phosphatase-1. Active fragments of inhibitor-1 showed little secondary or tertiary structure. The chemical shifts the inactive dephospho form and the active phospho form had only a localized change around the phosphorylated threonine. Phosphorylation of inhibitor-1 does not induce a major conformational change but instead provides a strong negative charge that is essential for binding to protein phosphatase-1.

Structural thermodynamic calculations of an ensemble of microcystin-LR and microcystin-LL free solution structures docked onto protein phosphatase-1 indicate that

many different conformations and positions are available for these toxins to bind at a similar free energy of disassociation/binding. However, these free energies of disassociation/binding have different entropy and enthalpy contributions highlighting the entropy-enthalpy compensation effect. Binding of microcystins is a dynamic process despite no major conformational changes, and is generally entropy driven.

Acknowledgements

I would like to thank my supervisor Dr. Brian Sykes for his guidance and patience. He provided an exceptional laboratory to work in that had excellent people and was well equipped. Considering that I had no previous NMR experience before my graduate studies I can safely say that I have learned many new ideas and concepts during this degree. It was also fun to play golf and go out for Friday lunches when I was able to.

I would also like to thank my "unofficial co-supervisor" Dr. Charles Holmes. He was always very positive, and was excited about the structural possibilities of protein phosphatases and their inhibitors. I found his energy and enthusiasm for the research to be enriching.

There are three people who made outstanding contributions to my thesis. They were co-authors on the papers I submitted or published and I learned much from their detailed knowledge and experience. The first is Pierre Lavigne. He was the first author on the Structural Thermodynamic Calculation paper that was submitted to Protein Science (first section of chapter 4). He made major contributions to the writing, figure preparation, research work, and the development of STC. He is an expert in thermodynamics and I learned a lot from him considering I had only basic knowledge before entering graduate studies. He was also extremely patient and somebody who I could talk to.

The second is Frank Sönnichsen who initially provided me with strong guidance and knowledge of NMR. He was instrumental in helping me to determine my first solution structure, microcystin-LR, using NMR. Despite being only a peptide, it was challenging because of the uncommon and D-amino acids. We had to work through a lot of problems associated with using the Insight II software.

The third is Robert Boyko who was the computer programmer behind the construction of STC. It was a pleasure to work with him in designing this program. I explained to him everything that I had to do myself when doing structural thermodynamic calculations, then he was able to automate it. Thanks to him instead of taking a day to perform calculations on one complex, we where able to do 60 complexes in a day using STC. He always was very "positive" towards me as well.

There are many people who made significant contributions to my research. From the Sykes Lab, Stéphane Gagné was a helpful person to go to when I experienced technical problems. Leigh Willard was involved in helping Pierre and I with STC. She was a very nice person to be around. Pascal Mercier was very helpful at the end of my thesis when I was writing up. His knowledge of Macs and inserting figures into Word was invaluable. Pascal is a very outgoing person. Gerry McQuaid is an expert at maintaining the magnets and was also helpful when I was first gaining hands-on experience in running the spectrometers. Other lab members that were nice to be around were Linda Saltibus who was very helpful and nice to talk to, Monica Li who always had a smile, Leo Spyracopoulos who was helpful when I had technical problems and always tried to make me laugh at least once every five minutes, Carolyn Slupsky who helped me with NMR details, and Dave Corson who always had something to say. People who had come and gone but were helpful were Wendy Findlay, Larry Calhoun, Tim Jellard and Gillian Henry who gave me a nice Red Dwarf shirt. Other Sykes Lab members that were nice to be around where Campbell McInnes, Key-Sun Kim, Jian-Jun Wang, Bruce Lix, Patricia Campbell, Krishna Rajarathnam, Matthew Crump, and David Wishart.

From the Holmes Lab, there was John Dawson who worked in a collaboration comparing protein phosphatase-1 with calcineurin which resulted in a publication that has been submitted based on the work. Hue Anh Luu was nice and helpful in the expression and purification of protein phosphatase-1. Marcia Craig helped me in purifying microcystins. Tara McCready was involved in the inhibitor-1 work and some of the NMR data ended up in a paper that we were co-authors on.

There were others who make significant contributions to my activities. The first is Sue Smith who helped with the scholarship grants and other business that I had to attend to. An excellent person to talk to and always tried to make me feel better. Paul Semchuk did an excellent job for all the peptide synthesis work in making inhibitor-1.

I would also like to acknowledge NSERC and the Alberta Heritage Foundation for Medical Research for providing financial support through my degree.

I would like to thank my parents, John and Marilynn, for their support of my academic activities despite difficulties from various sources. They did their best in raising me and I will always appreciate it.

Finally, I would like to thank Alix, the very special person in my life. I do not know what life would be like without her.

Table of Contents

Chapter 1- Introduction 1
Part 1- Description of Ser/Thr Protein Phosphatases 2
1. Protein Phosphatase-1 3
2. Protein Phosphatase-2A 5
3. Protein Phosphatase-2B (Calcineurin) 5
4. Protein Phosphatase-2C 7
Part 2- Physiological Roles for Serine/Threonine Protein Phosphatases 8
1. Glycogen Metabolism 8
2. Muscle Contractility 9
3. Protein Synthesis 10
Part 3- The Toxins and Inhibitors of Serine/Threonine Phosphatases 12
1. Microcystins (e.g. Microcystin-LR) 12
2. Nodularins (e.g. Motuporin) 14
3. Okadaic Acid 15
4. Calyculins (e.g. Calyculin A) 15
5. Inhibitor-1 16
Part 4- Objectives and Goals of Thesis 17
References 20

Chapter 2- A comparison of nuclear magnetic resonance solution structures of microcystin-LR and motuporin 24
Summary 25
Introduction 26
Methods 28
Results and Discussion 29
Chemical Shift Analysis of Microcystin-LR at 5°C 29
Solution Structures of the Peptides 31
Comparison of Solution Structures to Modeled Structures 34
Comparison of Microcystin-LR and Motuporin Solution Structures 35
References 37

Chapter 3- A molecular basis for different interactions of PP-1c toxins: molecular models for bound motuporin, okadaic acid, calyculin A, and microcystin 39
Summary 40
Introduction 41
Methods 45
Results 48
Comparison of Free NMR Microcystin-LR With Bound Crystal Microcystin-LR 48
Superimposition of Free NMR Motuporin Onto Crystal PP-1c 50
Superimposition of Free Crystal Structures of OA and Calyculin A onto bound PP-1c 53
Determination of Predicted Molecular Surfaces of Free and Bound Marine Toxins 55
Discussion 59
References 65

Chapter 4- Use of structure-based thermodynamics in the analysis of protein phosphatase-1 and microcystin-LR and -LL docked complexes 68
Summary 69
Introduction 70
Methods 73
Docking Procedure 73
Structure-based Free Energy Calculations 73
Calculation of Enthalpy 74
Calculation of Entropy 75
The STC Program Suite 77
Results and Discussion 79
Calculation of the thermodynamics of dissociation of MLR and PP-1c from the X-ray complex 79
Assessment of the Docked Complexes Obtain From an Ensemble of NMR Solution Structures for MLR 85
Enthalpy-Entropy Compensation Phenomenon Evidenced From Structure-based Thermodynamics 92
Cooperativity of Dissociation and Multiple Conformations (Microstates) for MLR:PP-1c Complex 97
Implications for Molecular Recognition 101
Analysis of NMR Docked MLR:PP-1c Complexes: Part 1- All Complexes 103
Analysis of NMR Docked MLR:PP-1c Complexes: Part 2- 32 High Affinity Complexes 109
Analysis of NMR Docked MLR:PP-1c Complexes: Part 3- 16 Low Affinity Complexes 119
Analysis of NMR Docked MLR:PP-1c Complexes: Part 4- 7 Most Probable Complexes 126
Analysis of NMR Docked MLL:PP-1c Complexes: Part 1- All Complexes 136
Analysis of NMR Docked MLL:PP-1c Complexes: Part 2- 37 High Affinity Complexes 137
Analysis of NMR Docked MLL:PP-1c Complexes: Part 3- 20 Low Affinity Complexes 147
Analysis of NMR Docked MLL:PP-1c Complexes: Part 4- 6 Best Complexes 157
References 158

Chapter 5- Structural comparison of protein phosphatase-1 with calcineurin: explanation for why microcystin is unable to inhibit calcineurin 164
Summary 165
Introduction 166
Methods 168
Results and Discussion 169
Comparison of the Crystal Structures of Protein Phosphatase-1c and Calcineurin A 169
References 182

Chapter 6- Structural comparison of the inactive dephospho form and active phospho form of inhibitor-1 183
Summary 184
Introduction 185
Methods 186
Results and Discussion 187
Chemical Shift Analysis of Inactive Dephospho and Active Phospho Inhibitor-1 Fragments 187
Comparison of 9-54 Fragments With 1-54 fragments of Inhibitor-1 193
Comparison Between Inactive Dephospho Inhibitor-1 and Active Phospho Inhibitor-1 195
Use of CSI to Analyze the Chemical Shifts of Inhibitor-1 197
Coupling Constant Analysis of Inhibitor-1 1-54 198
Comparison to Previously Studied Dephospho verse Phospho Proteins 199
Implications for Bound Form of Inhibitor-1 200
References 202

Chapter 7- $^{15}N/^{13}C$ Labeled microcystin-LR bound to protein phosphatase-1c 204
Summary 205
Introduction 206
Methods 207
Results and Discussion 211
References 220

Chapter 8- Conclusion 208
Part 1- Overview of Accomplishments 222
Part 2- Future Research Directions 225
References 228

Appendix 1- Using STC 229

List of Tables

Chapter 2

Table 1- Proton Chemical Shift Analysis of Microcystin-LR at 5°C ... 29
Table 2- Proton Chemical Shift Analysis of Microcystin-LR at 25°C ... 30

Chapter 3

Table 1- RMSD Between Starting and Final Docked Structures for Each Toxin ... 47

Chapter 4

Table 1- Thermodynamic Parameters for MLR:PP-1c Complex ... 84
Table 2- Overview STC results ... 156

Chapter 5

Table 1- Sequential Alignment of Protein Phosphatase-1, -2A, and Calcineurin ... 167
Table 2- PP-1c and Corresponding Calcineurin Residues in the MLR Binding Site ... 173
Table 3- Identical Residues in Calcineurin A and PP-1c at the MLR Binding Site ... 174
Table 4- Non-conserved Residues When Comparing Calcineurin and PP-1c ... 174
Table 5- Structural Analysis of Differing Residues Between 2 Enzymes ... 176

Chapter 6

Table 1- Chemical Shifts of Dephospho Inhibitor-1 Residues 9-54 ... 187
Table 2- Chemical Shifts of Phospho Inhibitor-1 Residues 9-54 ... 189
Table 3- Chemical Shifts of Dephospho Inhibitor-1 Residues 1-54 ... 190
Table 4- Chemical Shifts of Phospho Inhibitor-1 Residues 1-54 ... 192
Table 5- Comparison of Fragments 9-54 with 1-54 for Dephospho Inhibitor-1 ... 194
Table 6- Comparison of Fragments 9-54 with 1-54 for Phospho Inhibitor-1 ... 195
Table 7- Chemical Shift Changes Due to Phosphorylation of Inhibitor-1 1-54 ... 196
Table 8- CSI analysis of HA Protons for Inhibitor-1 1-54 ... 197
Table 9- Coupling Constant Analysis of Inhibitor-1 1-54 ... 198

Chapter 7

Table 1- ^{1}H, ^{15}N, and ^{13}C Chemical Shift Assignments for Microcystin-LR ... 211

List of Figures

Chapter 1

Figure 1- General mechanism for reversible protein phosphorylation 2
Figure 2- General comparison of serine/threonine protein phosphatases 3
Figure 3- Diagram of the structure of protein phosphatase-1 4
Figure 4- Protein phosphatase-1 regulation of glycogen metabolism................................ 9
Figure 5- Role of protein phosphatase-1 in increasing protein synthesis........................ 10
Figure 6- Primary sequence of microcystins, nodularins, okadaic acid, and calyculin A 14

Chapter 2

Figure 1- Primary sequences of microcystin-LR and motuporin 26
Figure 2- Nomenclature for microcystin-LR atoms.. 31
Figure 3A- Ensemble of calculated NMR solution structures of microcystin-LR 32
Figure 3B- Ensemble of calculated NMR solution structures of motuporin.................... 33
Figure 4- Superimposition of microcystin-LR and motuporin average structures 34

Chapter 3

Figure 1- Primary sequences of microcystin, nodularin, okadaic acid, and calyculin A .. 42
Figure 2- Comparing free solution and bound crystal structures of microcystin-LR 48
Figure 3- Free solution structure of microcystin-LL 50
Figure 4- Comparison of bound microcystin-LR with motuporin................................... 51
Figure 5A- Crystal structure complex of microcystin-LR bound to PP-1c...................... 52
Figure 5B- Docked motuporin:PP-1c complex 52
Figure 6A- Overview of microcystin-LR:PP-1c complex 53
Figure 6B- Overview of motuporin:PP-1c complex 53
Figure 7A- Comparison of okadaic acid with microcystin-LR 54
Figure 7B- Docked model of okadaic acid:PP-1c complex 54
Figure 8A- Comparison of calyculin A with microcystin-LR... 55
Figure 8B- Docked model of calyculin A:PP-1c complex.. 55
Figure 9- Molecular surfaces of marine toxins using GRASP 56
Figure 10- Distance molecular surfaces of marine toxins using GRASP 57
Figure 11- Distance molecular surfaces of PP-1c using GRASP 58

Chapter 4

Figure 1- Comparison of bound with ensemble of free structures of microcystin-LR 80
Figure 2- GRASP image of MLR:PP-1c complex 81
Figure 3A- Changes in ASA per residue upon dissociation of crystal complex 82
Figure 3B- Corresponding free energy of dissociation changes per residue 83
Figure 4- Schematic representation of crystal and docked complexes 86
Figure 5- Free energy of dissociation verses positional RMSD 88
Figure 6- Free energy of dissociation verses free energy of the complex 90
Figure 7- Corrected free energy of dissociation verses free energy of complex 91
Figure 8- Corrected free energy of dissociation verses positional RMSD 92
Figure 9- Seven best microcystin:PP-1c complexes 93
Figure 10A- Corrected dissociation enthalpy verses corrected dissociation entropy 95
Figure 10B- Complex dissociation enthalpy verses complex dissociation entropy 96
Figure 11A- Probability of a complex verses positional heavy RMSD 99
Figure 11B- Probability of a complex verses positional backbone RMSD 100
Figure 12- Dissociation free energies of all MLR:PP-1c complexes 103
Figure 13- High affinity MLR:PP-1c complexes 104
Figure 14- Corrected dissociation free energies of all MLR:PP-1c complexes 105
Figure 15- Entropy and enthalpy contributions for all MLR:PP-1c complexes 106
Figure 16- Burial of ASA by microcystin-LR for every complex 107
Figure 17- Burial of ASA by PP-1c for every complex 108
Figure 18- Free energy of dissociation for high affinity MLR:PP-1c complexes 109
Figure 19- Corrected free energy for high affinity MLR:PP-1c complexes 110
Figure 20- Entropy and enthalpy contributions for high affinity MLR:PP-1c 111
Figure 21- Burial of ASA by microcystin-LR for high affinity complexes 112
Figure 22- Burial of ASA by PP-1c for high affinity complexes 113
Figure 23- Burial of ASA by Adda of MLR for high affinity complexes 114
Figure 24- Burial of ASA by Arg of MLR for high affinity complexes 115
Figure 25- Burial of ASA by Leu of MLR for high affinity complexes 116
Figure 26- Burial of ASA by backbone of MLR for high affinity complexes 117
Figure 27- Burial of ASA by microcystin-LR for low affinity complexes 119
Figure 28- Burial of ASA by PP-1c for low affinity complexes 120
Figure 29- Burial of ASA by Adda of MLR for low affinity complexes 121
Figure 30- Burial of ASA by Arg of MLR for low affinity complexes 122
Figure 31- Burial of ASA by Leu of MLR for low affinity complexes 123
Figure 32- Burial of ASA by backbone of MLR for low affinity complexes 125
Figure 33- Dissociation free energy for 7 most probable MLR:PP-1c complexes 126

Figure 34- Corrected free energy for 7 most probable MLR:PP-1c complexes............. 127
Figure 35- Enthalpy and entropy contributions for most probable complexes 128
Figure 36- Burial of ASA by microcystin-LR for most probable complexes 129
Figure 37- Burial of ASA by PP-1c for most probable complexes 130
Figure 38- Burial of ASA Adda of MLR for most probable complexes........................ 131
Figure 39- Burial of ASA Arg of MLR for most probable complexes.......................... 132
Figure 40- Burial of ASA Leu of MLR for most probable complexes.......................... 133
Figure 41- Burial of ASA backbone of MLR for most probable complexes 135
Figure 42- Free energy of dissociation for every MLL:PP-1c complex 136
Figure 43- Docked conformations for the 37 high affinity MLL:PP-1c complexes....... 137
Figure 44- Free energy of dissociation for high affinity MLL:PP-1c complexes 138
Figure 45- Enthalpy and entropy contributions for high affinity complexes 139
Figure 46- Burial of ASA by MLL for 37 high affinity MLL:PP-1c complexes 140
Figure 47- Burial of ASA by PP-1c for 37 high affinity MLL:PP-1c complexes.......... 141
Figure 48- Burial of ASA by Adda of MLL for high affinity MLL:PP-1c complexes .. 142
Figure 49- Burial of ASA by new Leu of MLL for high affinity complexes................. 143
Figure 50- Burial of ASA by Leu of MLL for high affinity MLL:PP-1c complexes..... 144
Figure 51- Burial of ASA by backbone of MLL for high affinity complexes 146
Figure 52- Docked conformations for the 20 low affinity MLL:PP-1c complexes........ 147
Figure 53- Free energy of dissociation for low affinity MLL:PP-1c complexes 148
Figure 54- Enthalpy and entropy contributions for low affinity complexes 149
Figure 55- Burial of ASA by MLL for 20 low affinity MLL:PP-1c complexes 150
Figure 56- Burial of ASA by PP-1c for 20 low affinity MLL:PP-1c complexes........... 151
Figure 57- Burial of ASA by Adda of MLL for low affinity MLL:PP-1c complexes ... 152
Figure 58- Burial of ASA by new Leu of MLL for low affinity complexes.................. 153
Figure 59- Burial of ASA by Leu of MLL for low affinity MLL:PP-1c complexes...... 154
Figure 60- Burial of ASA by backbone of MLL for low affinity complexes 156
Figure 61- Docked conformations for 6 best MLL:PP-1c complexes........................... 157

Chapter 5

Figure 1- Crystal structure complex of MLR bound to PP-1c 169
Figure 2- Comparison of PP-1c and calcineurin crystal complexes 170
Figure 3- Comparison of 8 conserved residues in MLR binding pocket of PP-1c 171
Figure 4- Attempted docking of MLR onto calcineurin 172
Figure 5- Grasp image of MLR positioned onto calcineurin 175
Figure 6- Most significant structural difference between PP-1c and calcineurin 177
Figure 7- Interaction with His 339 and Asp 313 of L7 loop in calcineurin 179
Figure 8- Overview of PP-1c and calcineurin showing His-Asp interaction 180
Figure 9- Explanation for why immunosuppressants do not bind to PP-1c 181

Chapter 7

Figure 1- ^{15}N-HSQC spectrum of free microcystin-LR 213
Figure 2- ^{15}N-HSQC spectrum of bound microcystin-LR 214
Figure 3- ^{13}C-HSQC spectrum of free microcystin-LR (0-80ppm) 215
Figure 4- ^{13}C-HSQC spectrum of bound microcystin-LR (0-80ppm) 216
Figure 5- ^{13}C-HSQC spectrum of free microcystin-LR (110-150ppm) 217
Figure 6- ^{13}C-HSQC spectrum of bound microcystin-LR (110-150ppm) 218

Appendix 1

Figure 1- STC main menu 229
Figure 2- Input of pdb file into STC 230
Figure 3- Calculation of thermodynamics 231
Figure 4- Results of ASA and thermodynamic calculations 232

Abbreviations

A	Average
Å	Angstrom
Adda	β-amino acid [2S, 3S, 8S, 9S]-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid
ASA	Accessible Surface Area
C	Crystal
CaM	Calmodulin
cAMP	cyclic AMP
CnB	Calcineurin B
comp	complex
corr	corrected
COSY	COrrelated SpectroscopY
ΔC_P	change in heat capacity
CSI	Chemical Shift Index
DMSO	DiMethyl SulfOxide
ds-RNA	double stranded RNA
eIF-2	eukaryotic Initiation Factor 2
ΔG_{bind}	change in free energy of binding
ΔG_d	change in free energy of dissociation
ΔH	change in enthalpy
HSQC	Heteronuclear correlation spectroscopy
hsp90	heat shock protein 90
I-1	Inhibitor-1
IC_{50}	Inhibitory Concentration at 50%
K	Kelvin
kcal	kilocalories
kDa	kiloDaltons
Masp	D-erythro-β-methyl aspartic acid
Mdha	N-methyldehydroalanine
Mdhb	N-methyldehydroalanine
MHz	MegaHertz
MLL	Microcystin-LL
MLR	Microcystin-LR
mol	mole
NF-AT	Nuclear Factor of Activated Transcription

NMR	Nuclear Magnetic Resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
np	nonpolar
OA	Okadaic Acid
PK	Protein Kinase
PKA	Protein Kinase A
pol	polar
PP-1	Protein Phosphatase-1
PP-1c	Protein Phosphatase-1 catalytic subunit
PP-2A	Protein Phosphatase-2A
PP-2Ac	Protein Phosphatase-2A catalytic subunit
PP-2B	Protein Phosphatase-2B
PP-2C	Protein Phosphatase-2C
RMDS/rmsd	Root Mean Square Deviation
ΔS	change in entropy
ΔS_{bb}	change in conformation entropy backbone
$\Delta S_{bu \to ex}$	gain in conformational entropy in exposing side-chains
ΔS_{conf}	change in conformation entropy
$\Delta S_{ex \to u}$	change in conformation entropy for side-chains (secondary structure)
ΔS_{rt}	change in rotational/translational entropy
ΔS_{sol}	change in solvation entropy
STC	Structural Thermodynamic Calculations
TFE	TriFluoroEthanol
TOCSY	Total Correlated SpectroscopY

Thesis Chapter 1

Introduction

Part 1- Description of Ser/Thr Protein Phosphatases

Reversible phosphorylation of proteins is a major control mechanism for biochemical events in cells. Phosphorylation of serine, threonine, and tyrosine residues control a wide variety of regulatory proteins. Protein phosphatases are signal transducing enzymes that dephosphorylate cellular phosphoproteins[1]. Approximately 30% of intracellular proteins are phosphoproteins and the combined genes for protein kinases and phosphatases amount to approximately 4% of the eukaryotic genome[2]. Protein kinases and phosphatases catalyze the opposing and dynamic activities of protein phosphorylation and dephosphorylation respectively (Figure 1). While there are a large number of different protein kinases that tend to be highly specialized in activity, there are only a small number of serine/threonine protein phosphatases having broad and overlapping substrate specificify that are responsible for all the dephosphorylation activity in a cell. Protein phosphatases can be divided into two groups based on activity: serine/threonine phosphatases and tyrosine phosphatase, they are structurally dissimilar. A further subdivision can be made in the serine/threonine phosphatase group into 4 types: protein phosphatase-1, protein phosphatase-2A, protein phosphatase-2B (calcineurin), and protein phosphatase-2C. All of these phosphatases have highly similar catalytic subunits with the exception of protein phosphatase-2C (Figure 2). Each phosphatase will now be reviewed on an individual basis.

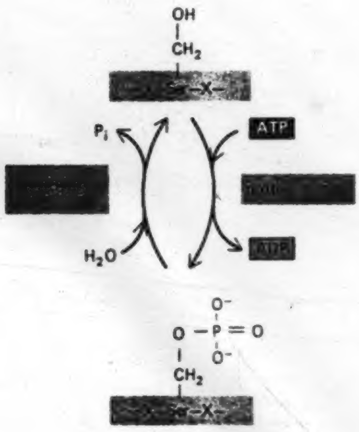


Figure 1-General mechanism for reversible protein phosphorylation. Protein kinases phosphorylate protein substrates, while protein phosphatases dephosphorylate protein substrates.

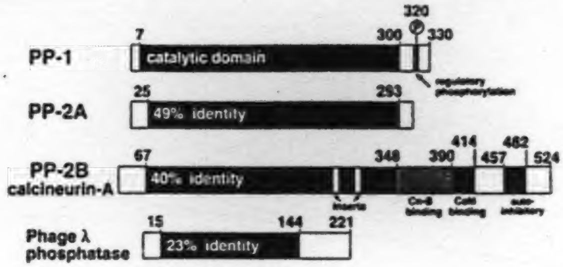


***Figure 2**-Diagram of the domain structure and regulatory regions of serine/threonine phosphatases[3]. Protein phosphatase-1 (PP-1) is from rabbit muscle while protein phosphatase-2A (PP-2A) and calcineurin A (PP-2B) are human. The black area represents the sequence covered in the protein phosphatase-1 crystal structure. Residues of calcineurin A implicated in the binding of calcineurin B (Cn-B), and calcium/calmodulin (CaM) are indicated by shading.*

1. Protein Phosphatase-1

Protein phosphatase-1 is distinguished from type 2 protein phosphatases (protein phosphatase -2A, -2B, -2C) on the basis of its ability to dephosphorylate the β-subunit of phosphorylase kinase and its inhibition by regulatory proteins inhibitor-1 and inhibitor-2 at nanomolar concentrations[4]. Type 2 protein phosphatases are unaffected by the inhibitor proteins and preferentially dephosphorylate the α-subunit of phosphorylase kinase. Protein phosphatase-1 has a 37-kd monomeric catalytic subunit[5]. This catalytic subunit when active normally associates with other regulatory subunits in the cell. For example, in muscle cells the catalytic subunit of protein phosphatase-1 forms a 1:1 complex with the G subunit. The G subunit is a 161 kDa protein that is responsible for binding to glycogen. Another subunit, M, complexes with the catalytic subunit of protein phosphatase-1 in myofibrils and is approximately 110 kDa in size. The physiological roles of these subunits will be discussed later.

The crystal structure of the 37 kDa catalytic subunit of recombinant mammalian protein phosphatase-1 was first determined in a complex with microcystin-LR[3]. The overall architecture of the enzyme is in a compact ellipsoidal form having dimensions of approximately 50 Å X 35 Å X 35 Å. Conceptually, the structure can be divided into two regions (Figure 3). The region from residues 7-182 is highly compact, and has a β–α–β–α–β metal coordinating unit which has two metal cations (e.g. Mn^{+2}) embedded in it. These cations are necessary for activity in the recombinant form but are not required for activity for native protein phosphatase-1. The region from residues 183-300 is a less compact irregular structure made up of mostly β-strands and contains the binding sites for the substrate and inhibitors. The crystal structure of microcystin bound to protein phosphatase-1c[3] was useful in comparing to the free NMR structure of microcystin and as a structure for docking solutions structures and crystal structures of various marine toxins (Chapters 3 and 4). Following this the crystal structure of human recombinant protein phosphatase-1 in the free form with no inhibitors bound was determined to have essentially the same structure as the bound form of protein phosphatase-1[6].

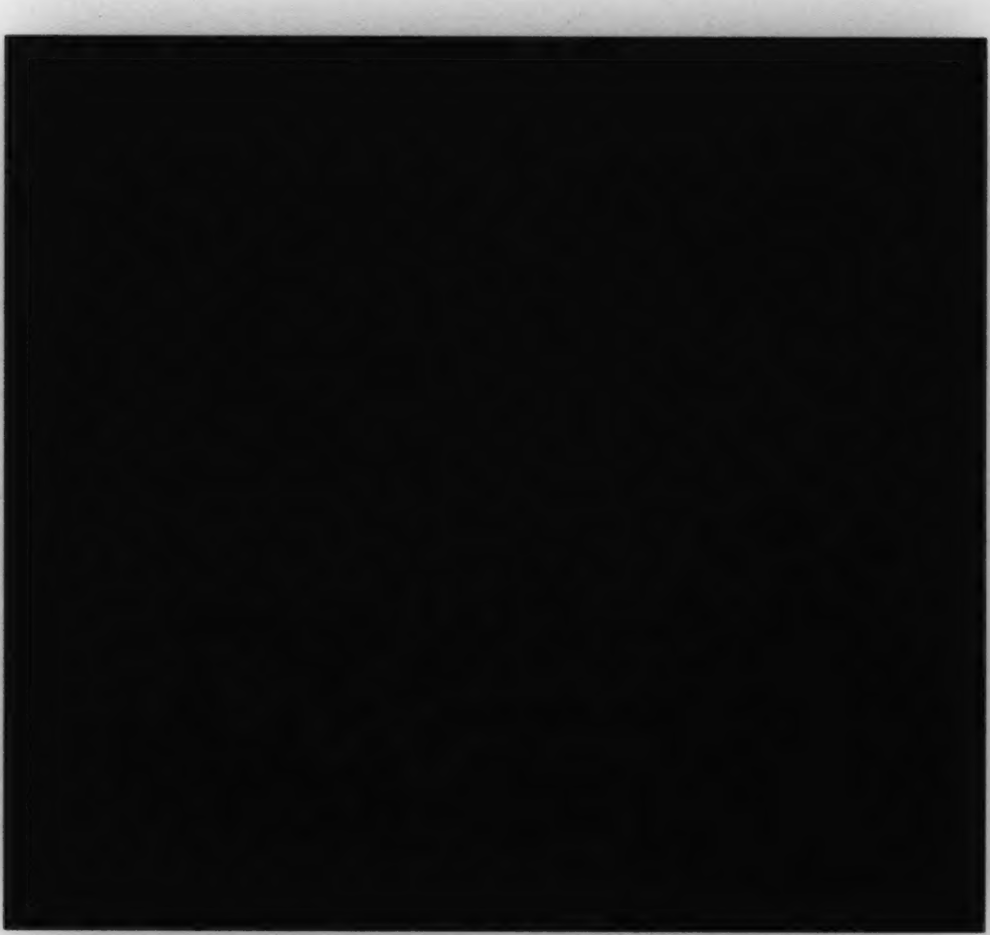


Figure 3-Crystal structure complex of MLR bound to protein phosphatase-1[3].

2. Protein Phosphatase-2A

Protein phosphatase-2A is one of three type 2 protein phosphatases that are distinguished from each other by their differing requirements for cations (if any). Protein phosphatase-2A like protein phosphatase-1 in the native form has no absolute requirement for bivalent cations for activity. However, protein phosphatase-2B (calcineurin) is dependent upon calcium (and also calmodulin) while protein phosphatase-2C requires magnesium[4].

The catalytic subunit of protein phosphatase-2A is highly similar to the catalytic subunit of protein phosphatase-1 (Figure 2) but is slightly smaller at 36 kDa[5]. Homology extends from residues 23 to 292 in protein phosphatase-2A with only a single deletion in protein phosphatase-1 and another in protein phosphatase-2A required to maximize homology. The overall sequence homology is 50% in residues 23-292, and rises to 67% if conservative substitutions are included. Protein phosphatase-2A has been identified in erythrocytes, skeletal muscle, heart muscle, nuclei associated with chromatin, and kidney mitochondria to name a few locations.

Native protein phosphatase-2A enzymes are heterotrimers of two regulatory subunits (A and B) and the catalytic subunit[7]. There is enough variation in the A and B regulatory subunits so that if the catalytic subunit could complex with each potential form of A and B subunits then over 40 different protein phosphatase-2A complexes could exist in mammalian cells. The role of the regulatory subunits A and B is to differentially inhibit the catalytic subunit activity with respect to specific substrates. Therefore, the regulatory subunits define the substrate specificity of the individual protein phosphatase-2A complex. Removal of the A subunit from a complex has been shown to alter substrate specificity.

3. Protein Phosphatase-2B (Calcineurin)

Protein phosphatase-2B (calcineurin) was first identified by its activity of dephosphorylating the α-subunit of phosphorylase kinase[5]. Later it was shown to be a calcium dependent calmodulin-stimulated enzyme. Calcineurin comprises approximately 1% of total brain protein yet has only 0.03% representation in skeletal muscle[5].

The brain form of calcineurin is a heterodimer consisting of a 61 kDa subunit which is known as calcineurin A and a smaller calcineurin B subunit of 19 kDa[5]. Calcineurin A is the catalytic subunit containing the calmodulin-binding site (Figure 2).

The C-terminus of the catalytic subunit is extended relative to the other serine/threonine protein phosphatases. This extended region contains a calcineurin B binding region, a calmodulin binding area, and an autoinhibitory site (Figure 2). Calcineurin B is the calcium binding regulating subunit and as such has four "EF hand" calcium binding domains that are four high affinity calcium binding sites in the native enzyme. This calcium subunit has homology with other members of the "EF hand" calcium binding proteins at the regions of the calcium binding loops. There are distinct differences though outside of these loops as indicated by sequence comparisons with calmodulin and troponin C being at 35% and 29% homologous overall sequence identity respectively. Another important feature of calcineurin B is the myristylation of the amino terminal glycine residue that may explain why it is associated with particulate fractions of the brain as well as the cytosol. It is unknown if this myristylation occurs in skeletal muscle where in it is soluble.

One of the most well known discoveries in the protein phosphatase field is the interaction that calcineurin has with immunosuppressive drugs that act by inhibiting phosphatase activity[7]. The target for these drugs is calcineurin. Inhibition of the dephosphorylation activity of this enzyme has the pharmocological effect of suppressing T cell activation and proliferation occur. Cyclosporin A, FK506, and rapamycin are common immunosuppressant drugs. In order to work they have to bind to an intracellular immunophilin receptor (e.g. cyclosporin A binds to cyclophilin and FK506 binds to FKBP12). This drug-receptor complex then binds to the calcineurin complex (subunits A and B) which mediates calcineurin inhibition. The consequence is to suppress T cell proliferation possibly by inhibiting interleukin-2 production via suppression of the nuclear factor of activated transcription (NF-AT) which is a transcription factor normally acting as a promoter of the interleukin-2 gene.

A ternary complex consisting of a calcineurin A fragment, calcineurin B, FKBP12 and FK506 was structurally solved by X-ray crystallography[8]. From this complex it was determined that the drug-immunophilin complex FK506 and FKBP12 does not contact the dephosphorylation catalytic site on calcineurin A. Instead of directly interfering with the catalytic site (more than 10 Å away from the active site), the FK506 and FKBP12 complex disrupts calcineurin activity by positioning itself to physically hinder the approach of a substrate to the active site. The crystal structure of calcineurin A in this complex was useful in a structural comparison to protein phosphatase-1 to explain why microcystin and other toxins are unable to inhibit calcineurin while actively inhibiting the similar protein phosphatase-1c (Chapter 5).

4. Protein Phosphatase-2C

Protein phosphatase-2C is a monomeric protein of 43 kDa and is magnesium dependent. This phosphatase is the most structurally distinct from the other serine/threonine protein phosphatases[7]. It is also the least studied of the serine/threonine protein phosphatases. *In vitro* it has high activity towards enzymes of cholesterol metabolism. Protein phosphatase-2C has resistance to marine toxins, being insensitive to such toxins like okadaic acid. In this manner it is similar to calcineurin but is different from protein phosphatase-1 and protein phosphatase-2A.

Part 2- Physiological Roles for Serine/Threonine Protein Phosphatases

Since serine/threonine protein phosphatases have multiple substrates that control a variety of cellular processes there are many different physiological roles for these enzymes. This section will focus on the physiological roles of protein phosphatases-1, 2A, and 2C. Specifically, glycogen metabolism, muscle contractility, and protein synthesis will be covered. Such examples illustrate the diverse mechanisms where protein phosphatases are key points of control. A major physiological role of calcineurin has already been examined in terms of its responsibility in controlling T cell activation and proliferation, and the effect of immunosuppressant drugs on this enzyme.

1. Glycogen Metabolism

Neuronal and hormonal control of glycogen metabolism is mediated via reversible protein phosphorylation of glycogen phosphorylase, phosphorylase kinase, and glycogen synthase[5]. Dephosphorylation by protein phosphatases results in inhibition of glycogen phosphorylase and phosphorylase kinase and activation of glycogen synthase. Protein phosphatase-1 and -2A (especially protein phosphatase-1) are the only enzymes in skeletal muscle that significantly dephosphorylate glycogen phosphorylase and glycogen synthase. Therefore, glycogen phosphorylase which converts glycogen to glucose when energy is required is deactivated into phosphorylase b (active form is phosphorylase a) by protein phosphatase-1 which dephosphorylates ser-14.

Indirect regulation of protein phosphatase-1 by the secondary messengers calcium and cAMP further illustrates the important role this enzyme has in glycogen metabolism[4]. Figure 4 illustrates the key role of protein phosphatase-1 in this process and also highlights the roles of protein phosphatase-2A, calcineurin, and inhibitor-1 in interacting with protein phosphatase-1. Adrenalin via cAMP results in the phosphorylation of two serines on the G subunit of protein phosphatase-1 via protein kinase A (PKA). Once this occurs the G subunit dissociates from the catalytic subunit of protein phosphatase-1. This dissociation of the two subunits of protein phosphatase-1 is the key event in these cAMP and calcium pathways. Once this occurs the catalytic subunit is free to interact with the glycogen metabolizing enzymes glycogen phosphorylase and glycogen synthase inhibiting glucose synthesis and activating glycogen synthesis respectively. The cAMP pathway can be attenuated by calcium which interacts with calcineurin and indirectly with protein phosphatase-2A causing dephosphorylation of different residues on the G subunit ironically also causing it to dissociate from the catalytic subunit of protein phosphatase-1 (phosphorylation by PKA at different sites also caused dissociation).

Dissociation of the G subunit from the catalytic subunit doesn't guarantee activity of protein phosphatase-1. If inhibitor-1 is phosphorylated then it actively inhibits protein phosphatase-1 catalytic subunit, essentially turning off the effects of cAMP and calcium. This too can be overcome though by inactivating inhibitor-1 by dephosphorylating Thr-35 which is accomplished by calcineurin.

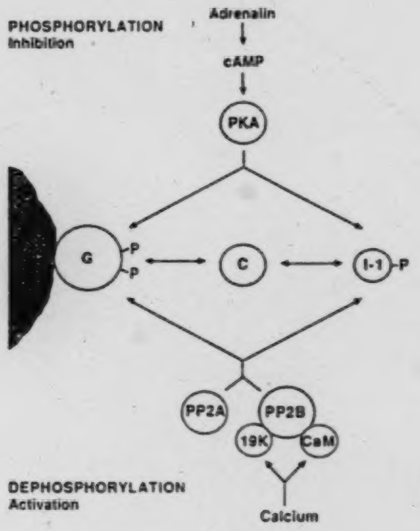


Figure 4*-Scheme summarizing the mechanisms by which cyclic AMP (cAMP) and calcium can regulate glycogen metabolism in skeletal muscle via protein phosphatase-1 catalytic subunit C and the G subunit*[4]*. Inhibitor-1 (I-1) can prevent the free catalytic subunit activity even after being freed from the G subunit. Calcineurin can both dissociate the G and C subunits of protein phosphatase-1 from a calcium stimulus and inactivate inhibitor-1. PKA is protein kinase A, while PP2A is protein phosphatase-2A.*

2. Muscle Contractility

The myosin P-light chain is phosphorylated by calcium/calmodulin-dependent myosin light chain kinases in striated and smooth muscles[9]. In smooth muscle this phosphorylation initiates contraction by causing myosin to interact with actin. This can be reversed when dephosphorylation of myosin occurs via protein phosphatase-1 happens. In skeletal and cardiac muscle, the myofibril-associated form of protein phosphatase-1 (specifically the M and catalytic subunits of protein phosphatase-1) is the enzyme that dephosphorylates the myosin P-light chain and thereby inhibit stimulation of the fibers.

3. Protein Synthesis

The overall rate of protein synthesis is enhanced by the dephosphorylation of eukaryotic initiation factor 2 (eIF-2) and by the phosphorylation of ribosomal protein S6[9]. The initiation factor eIF-2 in its dephosphorylated form is essential for positioning the initiator-tRNA on the small ribosomal 40S subunit. This function is blocked by phosphorylation of two serine residues on the α-subunit of eIF-2. The majority of dephosphorylation of eIF-2 is achieved by protein phosphatase-1 which is directly responsible activating this enzyme (Figure 5). In addition to directly interacting with eIF-2, protein phosphatase-1 also indirectly increases the dephosphorylation form by shutting off other proteins and kinases that are responsible for phosphorylating eIF-2 namely heat shock protein 90/heme-sensitive protein kinase and double-stranded-RNA-dependent protein kinase (Figure 5).

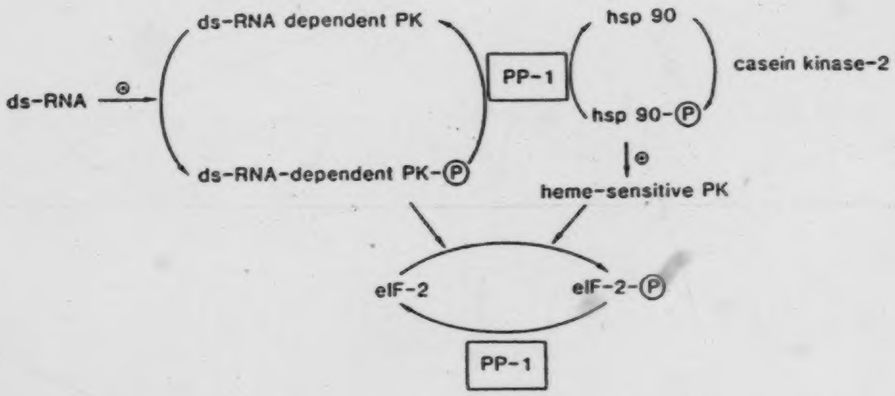


Figure 5-Role of protein phosphatase-1 in promoting the dephosphorylation of eIF-2 causing an increase in protein synthesis[9]. Protein phosphatase-1 can act directly on eIF-2 by directly dephosphorylating the initiation factor or it can also act indirectly by turning off two kinases, double-stranded RNA-dependent protein kinase and/or heme-sensitive protein kinase, which increase the amount of dephosphorylated eIF-2. Ds-RNA: double-stranded RNA, PK: protein kinase, hsp90: heat-shock protein of 90 kDa.

Opposite to its role of increasing protein synthesis by dephosphorylating eIF-2 protein phosphatase-1 can actually decrease protein synthesis by dephosphorylating S6 ribosomal protein[9]. The phosphorylation of S6 on 5 carboxy-terminal serine residues allows the 40S ribosomal subunit to form initiation complexes more efficiently. Protein phosphatase-1 dephosphorylates this ribosomal protein causing less efficient formation of initiation complexes. In order to account for this paradox on how protein phosphatase-1 can both increase and decrease protein synthesis two plausible explanations are that there are two distinct species of protein phosphatase-1 acting on different proteins or that protein kinases are responsible for the ultimate control of activating or deactivating protein phosphatase-1 (via inhibitor-1).

Part 3- The toxins and inhibitors of Ser/Thr Phosphatases

Reversible protein phosphorylation is the basis for control of many diverse cellular processes as mentioned in Parts 1 and 2. The basis for this control is through the phosphorylation and dephosphorylation of regulatory proteins. These regulatory proteins are phosphorylated on their serine, threonine, and/or tyrosine residues. Protein phosphatases are responsible for the dephosphorylation of these residues. Protein phosphatases specific to serine and threonine residues are normally regulated by a number of cellular inhibitors. There are also many different environmental marine toxins that are able to inhibit serine/threonine phosphatases. Such inhibition by marine toxins is outside the normal regulation of these cellular enzymes. These aberrant changes in the activities of these serine/threonine phosphatases lead to a number of diseases including cancer. The significant marine toxins that inhibit serine/threonine protein phosphatases will now be reviewed as well as a normal regulatory inhibitor protein, inhibitor-1.

1. Microcystins (e.g. Microcystin-LR)

The microcystin class of peptide hepatotoxins are metabolites of cyanobacteria in the genera *Microcystis* and *Anabaena* that grow worldwide in fresh and brackish water[10]. These toxins are responsible for extensive wildlife fatalities such as cattle or salmon (e.g. Pacific Northeastern salmon disease). Adverse effects on human health have also been recognized in countries where drinking water supplies or kidney dialysis fluids contain microcystin. For example, at least 40 human fatalities and 68 people were hospitalized due to microcystin contamination of kidney dialysis fluid in Brazil in 1996 (Time magazine April 22, 1996).

It has been shown that microcystin-LR is a potent inhibitor of the catalytic subunits of protein phosphatase-1 and -2A (PP-1c and PP-2Ac) as well as a powerful tumor promoter[11-15]. Inhibition of these enzymes in the liver by microcystin-LR is apparently associated with hepatocyte deformation due to reorganization of microfilaments[16]. The deformation is postulated to cause shrinking of the capillary cells within the liver. Consequently, blood cells are able to escape through the capillary lining entering the liver tissue causing internal bleeding in the liver often ending up with fatal consequences (e.g. Brazil 1996). In addition to large quantities of microcystin contamination immediately damaging the liver a second long-term health consequence of microcystin-LR in smaller doses are associated with tumor promotion in the liver. Liver tumor promotion may be linked to the ability of this cyclic peptide to promote hyperphosphorylation of cytokeratins associated with morphological changes in rat hepatocytes[17].

There are 2 important features of microcystin binding to PP-1c. The first is the high affinity that microcystin has for this enzyme. The K_i is $4 \cdot 10^{-11}$M reported by Takai et al.[18]. The second binding feature is the presence of a covalent linkage between the N-methyldehydroalanine residue of microcystin and Cys-273 of PP-1c. The covalent linkage takes hours to form and is not necessary for inhibition of the enzyme. Once formed though inhibition becomes irreversible.

A unique chemical feature that characterizes the microcystin peptide hepatotoxins is the presence of the unusual C_{20} β-amino acid [2S, 3S, 8S, 9S]-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid (Adda) (Figure 6)[19]. Other uncommon amino acid residues are D-erythro-β-methylaspartic acid (Masp) and N-methyldehydroalanine (Mdha) (Figure 1). Additionally, the Ala residue is in a nonstandard D conformation (instead of L) and the Glu as well as being a D-amino acid has its side chain incorporated into the cyclic backbone.

To date over 40 different microcystins have be characterized[20]. Within the microcystin family, members differ from each other in the nature of two variable L-amino acids indicated by suffix letters (e.g. L=Leu, R=Arg), and in the presence or absence of methyl groups on the D-erythro-β-methylaspartic acid (Masp) and/or N-methyldehydroalanine (Mdha) residues (Figure 6). Despite these differences no loss of protein phosphatase inhibition or tumor promotion has been observed.

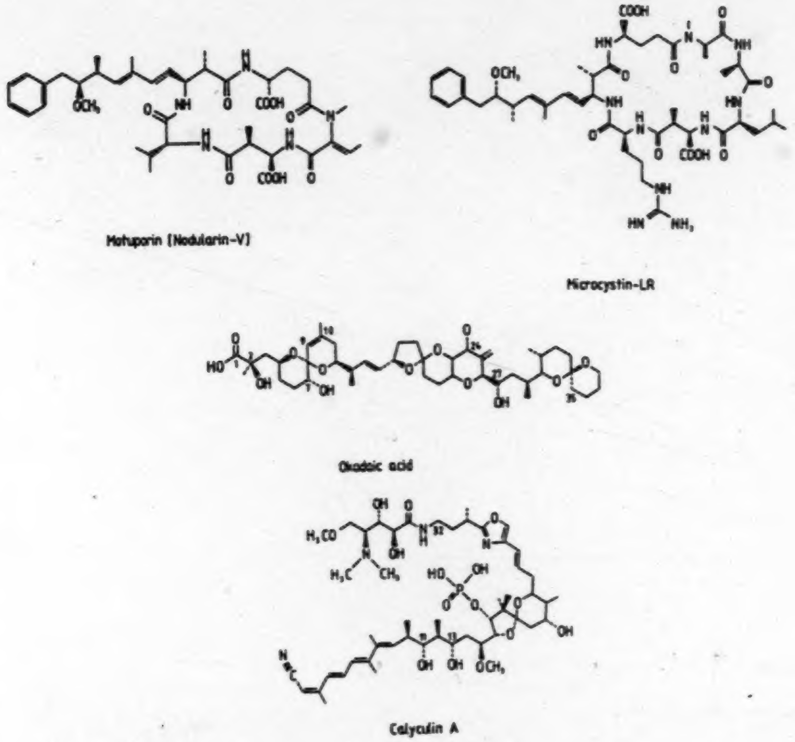


Figure 6- *Primary sequences of microcystin, nodularin, okadaic acid, and calyculin A. Residues for microcystin-LR (top,right) are β-linked D-erythro-β-methylaspartic acid (Masp); L-Arginine; β-[2S, 3S, 8S, 9S]-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid (Adda); γ-linked D-glutamic acid (Glu); N-methyldehydroalanine (Mdha); D-alanine; and L-Leucine. Motuporin aka. nodularin-V (top,left) has only 5 residues which are Masp; L-Valine; Adda; Glu; and N-methyldehydrobutyrine (Mdhb).*

2. Nodularins (e.g. Motuporin)

Nodularins are structurally related cyclic pentapeptides (Figure 6) that inhibit PP-1c and PP-2Ac with similar potency to microcystins and are also powerful tumor promoters[21]. The relatively hydrophobic cyclic pentapeptide motuporin (termed nodularin-V in Figure 6), was isolated from the marine sponge *T. swinhoei*[22]. Similar to microcystins, nodularins have the unusual Adda, Masp, and Glu residues (Glu side chain incorporated into the cyclic ring). Instead of having the Mdha residue as the microcystins do, nodularins have a modified version of this residue N-methyldehydrobutyrine (Mdhb). The difference between these two types of residues is the Mdhb has an addition methyl group attached to the beta-carbon of Mdha.

Although sharing similar biological properties in terms of liver toxicity and tumor promotion, important functional differences between the microcystins and nodularins have been identified. One difference is in the interaction with PP-1c and PP-2Ac. Although both toxins initially bind non-covalently and inhibit these enzymes, microcystin-LR, -LA, -LL undergo a secondary time-dependent interaction with the phosphatase[20,23]. This secondary interaction results in a covalent linkage causing irreversible modification of PP-1c/PP-2Ac. In contrast, nodularin or motuporin do not bind covalently to PP-1c/PP-2Ac after inhibiting it. A second difference is that in addition to acting as a tumor promoter, nodularin may also act as a carcinogen/tumor initiator[21].

3. Okadaic Acid

Okadaic acid and its analogs are polyether-like compounds (Figure 6), and as with microcystins and nodularins are potent inhibitors of PP-1c/PP-2Ac and powerful tumor promoters[16]. These toxins produced by the unicellular dinoflagellates *Dinophysis* and *Prorocentrum* are partly responsible for diarrhetic shellfish poisoning. Okadaic acid has also been used extensively as a research tool in studying reversible protein phosphorylation in many cellular processes[24].

4. Calyculins (e.g. Calyculin A)

Calyculin A (Figure 6) and its analogs isolated from the marine sponge *Discodermia calyx* are effective inhibitors of PP-1c and PP-2Ac, being tumor promoters as well[25]. Interestingly, the chemical structures of the microcystin/nodularin, okadaic acid, and calyculin A differ from each other (Figure 6) despite having the similar property of inhibiting PP-1c/PP-2Ac related to their tumor promotion activity.

5. Inhibitor-1

Inhibitor-1 is a 18.7 kDa thermostable protein first described[26] as a specific inhibitor for the catalytic subunit of protein phosphatase-1c (PP-1c)[27,28]. This inhibitor is one of many endogenous inhibitors controlling PP-1c activity in mammalian cells including DARPP-32[29], inhibitor-2[30], the ribosomal protein RIPP-1[31], the smooth muscle protein kinase C substrate C-kinase activated PP-1 inhibitor of Mr 17,000[32], and the nuclear protein NIPP-1[33] (reviewed in [34]).

Inhibitor-1 is active as a PP-1c inhibitor reducing its dephosphorylation activity only when it is phosphorylated by cAMP-dependent protein kinase at Thr-35[35]. Inhibitor-1 is regulated by reversible phosphorylation being dephosphorylated *in vivo* by protein phosphatase-2A and protein phosphatase-2B (calcineurin)[5,36,37]. Dephosphorylation of inhibitor-1 to its inactive form by calcineurin is a potential mechanism by which Ca^{2+} can increase PP-1c activity attenuating the effects of cAMP. Activation of PP-1c when inhibitor-1 is inactivated represents a potential phosphatase cascade, with inhibitor-1 lying at a critical junction of two secondary messenger systems[4].

An active fragment (residues 9-54) of inhibitor-1 was isolated[35] and found to retain full inhibitory activity against PP-1c if Thr-35 was phosphorylated. Proteolysis of this fragment into residues 22-54 and 13-41 showed that these fragments were inactive indicating functional importance in residues 9-22. This section contains many hydrophobic residues. It has been widely proposed that there are two independent structural elements in I-1 that are required for inhibition of PP-1c. These are the phosphorylation site at Thr-35 and residues 9-12 (KIQF)[29,38-40]. Models of I-1 and DARPP-32 bound to the crystal structure of PP-1c have proposed that the four sequential arginine residues preceding Thr-35 interact with acidic amino acids lining a groove located near the active site of PP-1c[3,39,41,42].

Part 4- Objectives and Goals of Thesis

The main hypothesis of this thesis is that the marine toxins microcystin, nodularin, okadaic acid, and calyculin A and the inhibitory protein inhibitor-1 all have common structural features that allow these ligands to inhibit protein phosphatase-1. To test this hypothesis the free and bound structures of the inhibitors would have to be determined, and then a comparison could be made between the different inhibitors. At the beginning of this thesis there was only three-dimensional structural information available for free okadaic acid and free calyculin A. There was no structural information available on microcystins, nodularins, and inhibitor-1, as well as for any serine/threonine protein phosphatase.

There are a variety of tools that I have utilized to determine or analyze protein structure. The first tool is nuclear magnetic resonance (NMR) which allows for the determination of free and bound solution structures. Free NMR solution structures of the ligands microcystin-LR, microcystin-LL, motuporin (nodularin-V), and inhibitor-1 (both inactive dephospho and active phospho fragments) are possible due to the availability and nature of the ligands. The other ligands okadaic acid[43] and calyculin A[44] have had their crystal structures determined. Determination of bound structures requires $^{15}N/^{13}C$ labeled inhibitors and milligram quantities of protein phosphatase-1 which can be achieved by expression and purification of a recombinant form. Due to the size of the catalytic subunit of protein phosphatase-1 at 37 kDa an entire thesis would have to be devoted to determining its structure via NMR.

A second tool is available if it is not possible to determine bound structures. Metropolis Monte Carlo docking (Insight II, Biosym Technologies) can be used to dock any free inhibitor structure onto protein phosphatase-1 if their structures have been determined. Fortunately, the crystal structure of protein phosphatase-1 (with microcystin-LR bound) was determined 3 years into my thesis[3] and the co-ordinates were made freely available to me prior to publication. This made it possible to construct bound models of any free ligand structure that had been determined allowing for a more complete and detailed view when comparing the toxin inhibitors of protein phosphatase-1.

A third tool known as structural thermodynamic calculations (STC) made it possible to do a detailed analysis of bound compounds. Basically, STC is able to analyze nonpolar and polar accessible surface area buried upon any ligand:enzyme complex formation. This information is then used to calculate the heat capacity of binding, the enthalpy of binding, the entropy of binding, the free energy of binding, and a dissociation constant of the complex. The measurements are theoritical being exclusively based on the coordinates of the three-dimensional ligand and enzyme structures. STC provides highly useful information about the bound complexes which aid in testing the main hypothesis of this thesis. STC is not limited to only crystal or NMR determined complexes but is also useful in studying modelled complexes. The ligand solution structures of the toxins determined by NMR are usually an ensemble of many calculated structures. Similar to an average solution structure, these calculated structures can be individually docked onto protein phosphatase-1. Having a large number of complexes to examine allows for a better analysis of the structure and function for the inhibitors of protein phosphatase-1.

A secondary hypothesis of this thesis is that there is some structural difference between protein phosphatase-1 and protein phosphatase-2B (calcineurin) that would provide an explanation for why the marine toxins and inhibitor-1 bind and inhibit protein phosphatase-1 while relatively being unable to inhibit calcineurin. This structural difference should be significant enough so that any mutagenesis studies could be used to replace protein phosphatase-1 residues by calcineurin residues causing a mutant form of protein phosphatase-1 that is able to resist being inhibited by the marine toxins. This hypothesis can be tested by comparing previously determined protein phosphatase-1[3] and calcineurin[8] crystal structures. Any structural differences that are identified can provide information about key residues involved in inhibitor binding for protein phosphatase-1 or prevention of inhibitor binding in the case of calcineurin.

This thesis is divided into 8 chapters including the Introduction and Conclusion (chapters 1 and 8 respectively). Chapter 2 deals with the determination of the solution structures for microcystin-LR and nodularin-V (motuporin). A comparison of these structures provides useful information about the similarities and differences in binding.

Chapter 3 is a comparison of the average free solution structure with the bound crystal structure of microcystin-LR. The structure of microcystin-LL is determined and compared with the more hydrophilic microcystin-LR. This chapter also contains information on the docked complexes for motuporin, okadaic acid, and calyculin A. An overview of the similarities between these diverse toxins in the free and bound (docked) forms is given.

Chapter 4 is the structural thermodynamic calculation (STC) chapter. This chapter contains a thermodynamic analysis of the bound crystal structure complex of microcystin-LR. A comparison is then made to other microcystin-LR and microcystin-LL complexes that had been generated by docking of the free solution structures of the toxins on an individual basis onto the crystal structure of protein phosphatase-1.

Chapter 5 is a comparison of the crystal structures of protein phosphatase-1 and calcineurin. An explanation for why microcystin-LR is able to inhibit protein phosphatase-1 while being unable to inhibit calcineurin is provided.

Chapter 6 is a structural analysis of inhibitor-1 based on chemical shifts. The chemical shifts of the inactive dephospho form are compared to the active phospho form of two different fragments that have similar activity to the complete protein. The reason why phosphorylation activates this inhibitor is given.

Chapter 7 is an analysis of the ^{15}N and ^{13}C spectra for free microcystin-LR and bound microcystin-LR. The difference in the free and bound spectra is pronounced and explained in this chapter.

References

1. Barford, D. (1995) *Current Opinion in Structural Biology* **5**, 728-734

2. Hunter, T. (1995) *Cell* **80**, 225-236

3. Goldberg, J., Huang, H.B., Kwon, Y.G., Greengard, P., Nairn, A.C. and Kuriyan, J. (1995) *Nature* **376**, 745-53

4. Cohen, P. and Cohen, P.T.W. (1989) *J Biol Chem* **36**, 21435-21438

5. Cohen, P. (1989) *Annu Rev Biochem* **58**, 453-508

6. Egloff, M.P., Cohen P.T.W., Reinemer, P. and Barford, D. (1995) *J Mol Biol* **254**, 942-959

7. Shenolikar, S. (1994) *Annu Rev Cell Biol* **10**, 55-86

8. Griffith, J.P., Kim, J.L., Kim, E.E., Sintchak, M.D., Thomson, J.A., Fitzgibbon, M.J., Fleming, M.A., Caron, P.R., Hsiao, K., and Navia, M.A. (1995) *Cell* **82**, 507-522

9. Bollen M. and Stalmans W. (1992) *Critical Reviews in Biochemistry and Molecular Biology* **27**, 227-281

10. Carmichael, W.W. (1994) *Scientific American* **270(1)**, 78-86

11. Honkanen, R.E., Zwiller, J., Moore, R.E., Daily, S.L., Khatra, B.S., Dukelow, M., and Boynton, A.L.(1990) *J Biol Chem* **265**, 19401-19404

12. Yoshizawa, S., Matsushima, R., Watanabe, M.F., Harada, K., Ichihara, A., Carmichael, W.W., and Fujiki, H. (1990) *J Cancer Res Clin Oncol* **116**, 609-614

13. MacKintosh, C., Beattie, K.A., Klumpp, S., Cohen, P., and Codd, G.A. (1990) *FEBS Lett* **264**, 187-192

14. Nishiwaki-Matsushima, R., Nishiwaki, S., Ohta, T., Yoshizawa, S., Suganuma, M., Harada, K., Watanabe, M.F., and Fujiki, H. (1991) *Jpn J Cancer Res* **82**, 993-996

15. Holmes, C.F.B. and Boland, M.P. (1994) *Current Opinion in Structural Biology* **3**, 934- 943

16. Eriksson, J.E., Toivola, D., Meriluoto, J.A.O., Codd, G.A., Kass, G.E.N., Karaki, J., Han, Y.G., and Hartshorne, D. (1990) *Biochem Biophys Res Commun* **173**, 1347-1353

17. Ohta, T., Nishiwaki, R., Yatsunami, J., Komori, A., Suganuma, M., and Fujuki, H. (1992) *Carcinogenesis* **13**, 2443-2447

18. Takai, A., Sasaki, K., Nagai, H., Mieskes, G., Isobe, M., Isono, K., and Yasumoto, T. (1995) *Biochem J* **306**, 657-665

19. Rhinehart, K.L., Harada, K.I., Namikoshi, M., Chen, C., Harvis, C.A., Munro, M.H.G., Blunt, J.W., Mulligan, P.E., Beasley, V.R., Dahlem, A.M., and Carmichael, W.W. (1988) *J Am Chem Soc* **110**, 8557-8558

20. Craig, M., McCready, T.L., Luu, H.A., Smillie, M.A., Dubord, P., and Holmes, C.F.B. (1993) *Toxicon* **31(12)**, 1541-1549

21. Ohta, T., Sueoka, E., Iida, N., Komori, A., Suganuma, M., Nishiwaki, R., Tatematsu, M., Kim, S., Carmichael, W.W., and Fujiki, H. (1994) *Cancer Research* **54**, 6402-6406

22. DeSilva, S.D., Williams, D.E., Andersen, R. J., Klix, H., Holmes, C.F.B., and Allen, T.M. (1992) *Tetrahedron Lett* **33**, 1561-1564

23. MacKintosh, R.W., Dalby, K.N., Campbell, D.G., Cohen, P.T.W., Cohen, P., and MacKintosh, C. (1995) *FEBS Letters* **371**, 236-240

24. Cohen, P., Holmes, C.F.B., and Tsukitani, Y. (1990) *Trends Biochem Sci* **15**, 98-102

25. Fujuki, H. and Suganuma, M. (1993) *Advances in Cancer Research* **61**, 143-151

26. Huang, F.L. and Glinsmann, W.H. (1976) *Eur J Biochem* **70**, 419-26

27. Cohen, P. (1978) *Curr Top Cell Reg* **14**, 117-196

28. Nimmo, G.A. and Cohen, P. (1978) *Eur J Biochem* **87**, 341-51

29. Hemmings, H.C., Nairn, A.C., Elliott, J.I. and Greegard, P. (1990) *J Biol Chem* **265**, 20369-76

30. Park, I.K. and DePaoli-Roach, A.A. (1994) *J Biol Chem* **269**, 28919-28

31. Beullens, M., Stalmans, W. and Bollen, M. (1996) *Eur J Biochem* **239**, 183-189

32. Eto, M., Ohmori, T., Suzuki, M., Furuya, K. and Morita, F. (1995) *J Biochem* **118**, 1104-7

33. Van Eynde, A., Wera, S., Beullens, M., Torrekens, S., Van Leuven, F., Stalmans, W. and Bollen, M. (1995) *J Biol Chem* **270**, 28068-74

34. Connor, J.H. et al. (1998) *J Biol Chem* **273**, 27716-24

35. Aitken, A., Bilham, T. and Cohen, P. (1982) *Eur J Biochem* **126**, 235-46

36. Hubbard, M.J. and Cohen, P. (1989) *Eur J Biochem* **186**, 711-720

37. Shenolikar, S. and Nairn, A.C. (1991) *Adv Secon Messenger Phosphoprotein Res* **23**, 1-121

38. Endo, S., Zhou, X., Connor, J., Wang, B. and Shenolikar, S. (1996) *Biochemistry* **35**, 5220-8

39. Kwon, Y.G., Huang, H.B., Desdouits, F., Girault, J.A., Greengard, P. and Nairn, A.C. (1997) *Proc Natl Acad Sci USA* **94**, 3536-41

40. Desdouits, F. et al. (1995) *Biochem Biophys Res Commun* **206**, 653-658

41. Connor, J.H., Quan, H., Oliver, C. and Shenolikar, S. (1998) *Methods Mol Biol* **93**, 41-58

42. Barford, D., Das, A.K. and Egloff, M.P. (1998) *Annual Reviews in Biophysical and Biomolecular Structure* **27**, 133-64

43. Tachibana, K., Scheuer, P.J., Tsukitani, Y., Kikuchi, H., Van Engen, D., Clardy, J., Gopichand, Y., Schmitz, F.J. (1981) *J Am Chem Soc* **103**, 2469-71

44. Kato, Y., Fusetani, N., Matsunaga, S., Hashimoto, K., Fujita, S., and Furuya, T.J. (1986) *J Am Chem Soc* **108**, 2780-1

Thesis Chapter 2

A comparison of nuclear magnetic resonance solution structures of microcystin-LR and motuporin

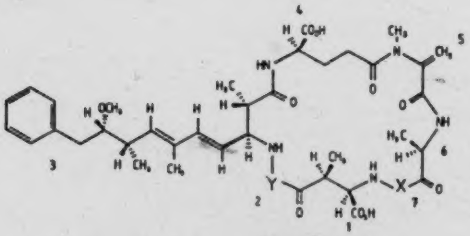
This chapter published as Bagu, J.R., Sönnichsen, F.D., Williams, D., Andersen, R.J., Sykes, B.D., and Holmes, C.F.B. (1995) *Nature Stuctural Biology* **2**, 114-116.

Summary

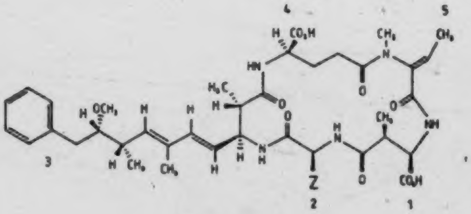
Hepatotoxic cyclic heptapeptide microcystins are powerful liver tumor promoters, potent inhibitors of protein phosphatase-1 and -2A, and are able to covalently bind to these enzymes. Hepatotoxic cyclic pentapeptide nodularins are also liver tumor promoters, potent inhibitors of protein phosphatase-1 and -2A, do not covalently interact with the protein phosphatases, and may additionally possess carcinogenic properties. Here we present the three dimensional structures of microcystin-LR and motuporin (nodularin-V) determined using ^{1}H Nuclear Magnetic Resonance spectroscopy. A comparison of the structures yields insights into how these cyanobacterial toxins might inhibit protein phosphatase-1 and -2A, why only microcystin-LR is able to covalently modify protein phosphatases, and provides a rationalization for why nodularins may also act as carcinogens.

Introduction

The microcystin and nodularin classes of peptide hepatotoxins are metabolites of cyanobacteria in the genera *Microcystis, Anabaena,* and *Nodularia* that grow worldwide in fresh and brackish water[1]. These toxins are responsible for extensive wildlife fatalities, and adverse effects on human health have also been recognized in countries where drinking water supplies contain cyanobacteria. Recently, it has been shown that microcystin-LR is a potent inhibitor of the catalytic subunits of protein phosphatase-1 and -2A (PP-1c and PP-2Ac) as well as a powerful tumor promoter[2-6]. Inhibition of these enzymes in the liver is apparently associated with hepatocyte deformation due to reorganization of microfilaments[7]. Liver tumor promotion may be linked to the ability of this cyclic peptide to promote hyperphosphorylation of cytokeratins associated with morphological changes in rat hepatocytes[8].



	AMINO ACID	
	X	Y
Microcystin-LR	Leu	Arg
Microcystin-LA	Leu	Ala
Microcystin-LL	Leu	Leu



	SIDE CHAIN
	Z
Nodularin-R	Arg
Nodularin-V	Val

Figure 1- *Primary sequences of microcystin-LR and motuporin. The residues for microcystin-LR (top structure) starting at 1. β-linked D-erythro-B-methylaspartic acid (Masp) and going clockwise are 2. L-arginine at Y, 3. β-[2S, 3S, 8S, 9S]-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid (Adda), 4. γ-linked D-glutamic acid (Glu), 5. N-methyldehydroalanine (Mdha), 6. D-alanine, and finally 7. L-leucine at X. Motuporin (nodularin-V, bottom structure) has only 5 residues instead of 7 which are 1. Masp, followed clockwise by 2. valine in place of arginine at Z, 3. Adda, 4. Glu, and 5. N-methyldehydrobutyrine (Mdhb, essentially Mdha with a methyl group added to the side chain).*

A unique chemical feature that characterizes the microcystin/nodularin peptide hepatotoxins is the presence of the unusual C_{20} β-amino acid [2S, 3S, 8S, 9S]-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid (Adda)[9]. To date over 40 different microcystins have be characterized[10]. Most of these microcystins, which are all cyclic heptapeptides, differ in the nature of two variable L-amino acids indicated by suffix letters (eg. L=Leu, R=Arg) and in the absence of methyl groups on the D-erythro-β-methyl aspartic acid (Masp) and/or N-methyldehydroalanine (Mdha) residues (Figure 1). Nodularins are structurally related cyclic pentapeptides (Figure 1) that inhibit PP-1c and PP-2Ac with similar potency to microcystins and are also powerful tumor promoters[11]. The relatively hydrophobic cyclic pentapeptide motuporin (also termed nodularin-V in Figure 1), which was recently isolated from the marine sponge *T. swinhoei*, is the only nodularin analog known[12].

Whilst sharing similar biological properties, important functional differences between the microcystins and nodularins have recently been identified. One difference is in the interaction with PP-1c. Although both toxins initially bind non-covalently and inhibit this enzyme, microcystin-LR undergoes a secondary time-dependent interaction with the phosphatase[13]. This secondary interaction results in an apparent covalent linkage causing irreversible modification of PP-1c. In contrast, nodularin or motuporin do not bind covalently to PP-1c after inhibiting it. A second difference is that in addition to acting as a tumor promoter, nodularin may also act as a carcinogen (tumor initiator) (Dr. H. Fujiki, personal communication).

Here we present the solution structures of microcystin-LR and motuporin determined using ^{1}H Nuclear Magnetic Resonance (NMR) spectroscopy. A comparison of the two structures yields insights into how they might both non-covalently bind and inhibit PP-1c and PP-2Ac, why microcystin-LR alone may then be able to covalently bind to PP-1c, and provides a possible explanation for why the nodularin class of tumor promoters may additionally act as carcinogens.

Methods

Peptides Preparation

Microcystin-LR and motuporin were purified from algal and sponge extracts respectively as previously described[10,12]. Samples used for NMR were 1-2mM dissolved in a 10mM potassium phosphate, 50mM sodium chloride buffer with 80-90% H_2O/10-20% D_2O at pH 7. 2,2-dimethyl-2-silapentane-5-sulfonic acid (0.1mM) was added as a ^1H NMR chemical shift standard.

Structure Determination and Analysis

All ^1H NMR spectra were recorded at 500MHz on a Varian VXR-500 NMR spectrometer. Proton NMR resonance assignments for microcystin-LR and motuporin were made using standard sequential assignment methods with double quantum filtered COSY, TOCSY, and NOESY two dimensional ^1H NMR spectra. ^1H-^1H internuclear distance restraints were obtained from two dimensional ^1H NOESY taken with a mixing time of 150ms so as to minimize spin diffusion. $^3J_{NH\alpha CH}$ and $^3J_{CH\text{-}CH}$ coupling constants were obtained by curvefitting one dimensional ^1H NMR spectra. Assignments were made at 25°C and 5°C, with structural information recorded at 5°C. For microcystin-LR, calibration of the NOEs was accomplished using the known distance of 1.8Å between the HB protons of the *sp*^{2} carbon of the Mdha side chain to scale the experimental crosspeak intensity. For motuporin, the averaged intensities of the crosspeaks from the two sets of backbone methylene protons of the Glu residue was scaled to a distance of 1.74Å. Sixty structures for each peptide were generated using the distance geometry program DGII in Biosym's InsightII version 2.2 (Biosym Technologies Inc.). For these cyclic peptides the vicinal angle restraints were very important in restraining the structure of the ring. Distance and angle violations were analyzed using the NMR refine module in InsightII. There were no consistently violated distance restraints over 0.1Å and angles fell within of the defined restraint range in most cases. For microcystin-LR, 46 structures converged to the same overall fold with the remaining 14 structures discarded due to violations. For motuporin, 48 structures converged and 12 were discarded because of violations. An average structure for each peptide was then generated from the remaining structures, and was subjected to constrained minimization to correct bond distances that had been distorted through averaging. The coordinates of the average structure with the corresponding calculated structures have been submitted to the Brookhaven Protein Data Bank.

Results and Discussion

Chemical Shift Analysis of Microcystin-LR and Motuporin

The proton chemical shifts of microcystin-LR at 5°C are given in Table 1. Table 2 has microcystin-LR chemical shifts at a different temperature of 25°C. Comparison of Table 1 with Table 2 indicates that the chemical shifts at both temperatures are almost identical with the exception of the amide protons which are temperature sensitive. This implies that the structure of microcystin-LR does not change between the temperatures of 5°C to 25°C. Figure 2 describes the nomenclature of the microcystin-LR atoms. This nomenclature was used so that Biosym InsightII correctly defined the atoms when calculating the solution structures. Motuporin has similar chemical shifts to microcystin-LR with the exception of a common valine instead of arginine and additional methyl group in the Mdhb residue verses the Mdha residue.

Table 1-Proton Chemical Shift Analysis of Microcystin-LR at 5°C

Atom	^{1}H Chemical Shift
Arg 4:HN	8.74
Leu 2:HN	8.46
Ala 1:HN	8.33
Glu 6:HN	8.23
Adda 5:HN	7.98
Masp 3:HN	7.81
Adda 5:HG	6.37
Mdha 7:HB1	5.94
Mdha 7:HB2	5.57
Adda 5:HE	5.55
Adda 5:HB	5.55
Adda 5:HA	4.48
Masp 3:HA	4.46
Ala 1:HA	4.41
Leu 2:HA	4.32
Arg 4:HA	4.26
Glu 6:HA	3.84
Adda 5:HH	3.49
Mdha 7:HM	3.42
Adda 5:H5	3.25
Masp 3:HB	3.22
Arg 4:HD	3.13
Adda 5:H1	3.04
Adda 5:H11	2.96
Glu 6:HG1	2.86
Adda 5:HF	2.77
Glu 6:HG2	2.56
Glu 6:HB	2.11
Leu 2:HB	2.01
Adda 5:H3	1.72
Leu 2:HG	1.63
Adda 5:H2	1.05
Adda 5:H4	1.03

Table 2-Proton Chemical Shift Analysis of Microcystin-LR at 25°C

Atom	^{1}H Chemical Shift
Arg 4:HN	8.58
Leu 2:HN	8.37
Ala 1:HN	8.21
Glu 6:HN	8.18
Adda 5:HN	7.94
Masp 3:HN	7.81
Adda 5:HM	7.37
Adda 5:HK	7.29
Adda 5:HG	6.33
Mdha 7:HB1	5.96
Adda 5:HB	5.59
Mdha 7:HB2	5.58
Adda 5:HE	5.55
Adda 5:HA	4.49
Ala 1:HA	4.46
Masp 3:HA	4.44
Leu 2:HA	4.33
Arg 4:HA	4.29
Glu 6:HA	3.91
Adda 5:HH	3.48
Mdha 7:HM	3.40
Adda 5:H5	3.26
Masp 3:HB	3.22
Arg 4:HD	3.14
Adda 5:H1	3.03
Adda 5:HI1	2.92
Glu 6:HG1	2.81
Adda 5:HF	2.75
Adda 5:HI2	2.74
Glu 6:HG2	2.58
Glu 6:HB1	2.03
Leu 2:HB1	1.99
Arg 4:HB1	1.96
Glu 6:HB2	1.85
Adda 5:H3	1.70
Leu 2:HG	1.65
Arg 4:HB2	1.62
Leu 2:HB2	1.53
Arg 4:HG	1.53
Masp 3:HG	1.38
Adda 5:H2	1.04
Adda 5:H4	1.01
Leu 2:HD1	0.90
Leu 2:HD2	0.86

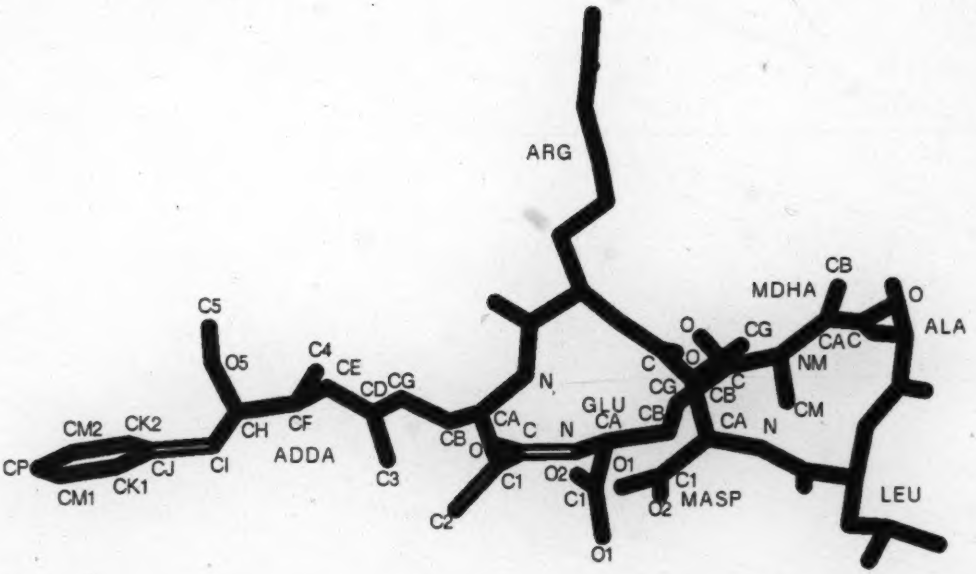


***Figure 2-** Nomenclature for atoms belonging to non-standard amino acids of microcystin-LR. This nomenclature was used in order for Biosym InsightII to correctly identify the atoms when calculating solution structures.*

Solution structures of the peptides

The family of solution structures of microcystin-LR determined by ^{1}H NMR is shown in Figure 3A. The root mean square deviation (RMSD) of the individual structures to the average structure based on the backbone atoms is 0.41Å +/- 0.10Å. The peptide ring can be described as having a saddle shaped motif with the carboxyl residues at the sides, Mdha residue front, and Arg at the back in the orientation shown. Both the Adda and Arg long side chains are highly flexible, while the Leu side chain is well defined. From the perspective of Figure 3A, the Adda side chain is located behind the saddle and the Arg side chain is above it. The structures were calculated from 87 NOE restraints (65 intraresidue, 17 sequential, and 5 long range) divided into 5 restraint classes: 1.8-2.7Å (18 NOEs), 1.8-3.3Å (14 NOEs), 1.8-3.9Å (13 NOEs), 1.8-4.5Å (3 NOEs), and 1.8-5.0Å (39 NOEs-29 of which were placed in this class because overlap or interference from zero quantum coherence made quantification difficult) and 4 NH-αCH vicinal angle restraints for the residues Masp ($^3J_{NH\alpha CH}$=9.8Hz), Arg ($^3J_{NH\alpha CH}$=9.5Hz), Adda ($^3J_{NH\alpha CH}$=9.6Hz), and Glu ($^3J_{NH\alpha CH}$=5.8Hz).

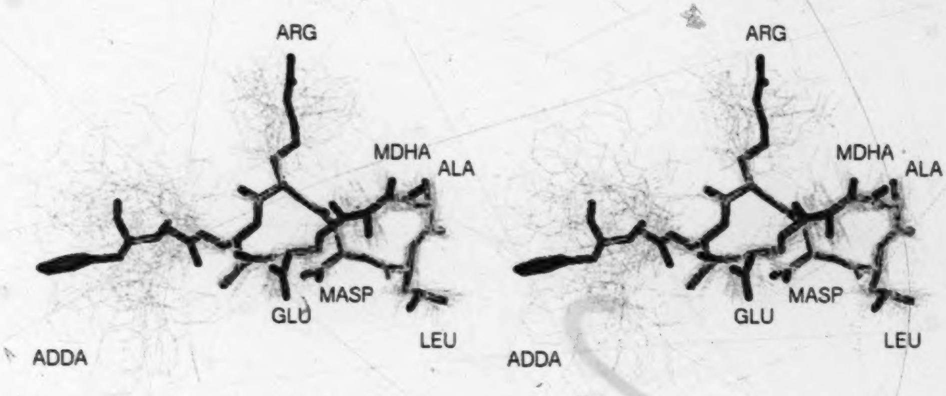


Figure 3A- *Ensemble of calculated NMR solution structures of microcystin-LR. The 46 calculated DGII structures are shown in grey; the thick black structure is the average structure. For the purpose of clarity hydrogen atoms are not shown. Backbone RMSD with respect to the average structure was 0.41Å +/- 0.10Å standard deviation.*

The family of solution structures of motuporin determined by ^{1}H NMR is shown in Figure 3B. The RMSD was 0.38Å +/- 0.08Å calculated for the backbone atoms. In a manner similar to microcystin-LR, the ring of motuporin forms a saddle shape, however it is smaller in size and less pronounced in pucker. Thus, the two share the same overall backbone structural motif. While the ring is well defined, the Adda side chain of motuporin is highly flexible. The Val in motuporin is less flexible than the Arg in microcystin-LR. In the orientation shown in Figure 3B the Adda side chain is behind the saddle. The structures were calculated from 46 NOE distance restraints (40 intraresidue, 5 sequential, and 1 long range) divided into 4 restraint classes: 1.8-2.7Å (12 NOEs), 1.8-3.3Å (8 NOEs), 1.8-3.9Å (5 NOEs), and 1.8-5.0Å (21 NOEs-18 of which were placed in this class because overlap or interference from zero quantum coherence made quantification difficult) and 2 NH-αCH vicinal angle restraints for residues Val ($^3J_{NH\alpha CH}$=9.9Hz) and Adda ($^3J_{NH\alpha CH}$=9.4Hz). The structures were checked against and found consistent with the αCH-backbone βCH vicinal angles for residues Masp ($^3J_{CH-CH}$=3.2Hz) and Adda ($^3J_{CH-CH}$=10.8Hz).

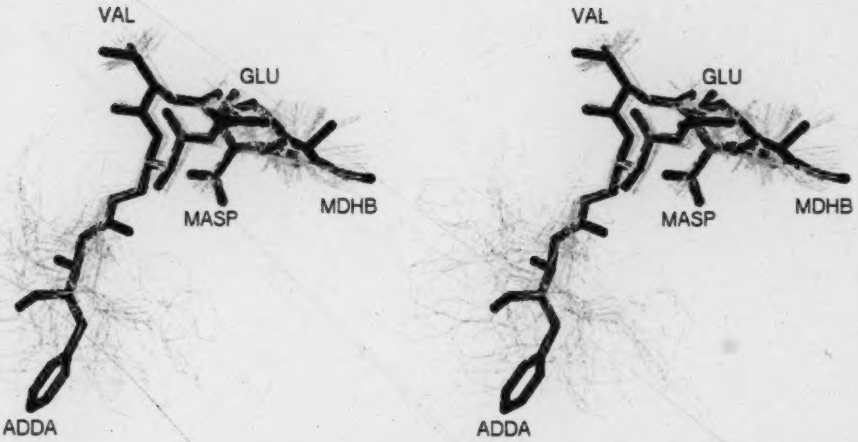


Figure 3B- *Ensemble of calculated NMR solution structures of motuporin. The 48 calculated DGII structures are shown in grey; the thick black structure is the average structure. Only heavy atoms are shown for clarity. The backbone RMSD to the average was 0.38Å +/- 0.08Å standard deviation.*

Figure 4 shows a superimposition of the average structures of microcystin-LR and motuporin. The backbone segment Masp-Arg-Adda in microcystin-LR was superimposed on the Masp-Val-Adda backbone atoms of motuporin (excluding the Adda carbonyl carbon). This superimposition was chosen because this range of atoms had less than 40 degrees difference between the two structures in measured psi, phi, and omega angles. There was no other significant range of atoms having less than 40 degree difference between the backbones in corresponding phi, psi, and omega angles. Structurally, the backbones of the two structures are very similar from the Masp carboxyl to the γ-linked D-glutamic acid (Glu) carboxyl through Arg/Val and Adda. In this superimposition, the positions of the Masp and Glu carboxyls in each peptide are similar. The replacement of Val in motuporin for Arg of microcystin-LR does not result in any major change in the position of the residue. The largest difference in the two structures can clearly be seen in the positioning of the Mdha in microcystin-LR verses N-methyldehydrobutyrine (Mdhb) in motuporin. The presence of the D-Ala and Leu residues in microcystin-LR, which are not in motuporin and have no replacements, result in a prominent upwards curvature in the Mdha, D-Ala, and Leu residues of the saddle in microcystin-LR that does not occur in motuporin.

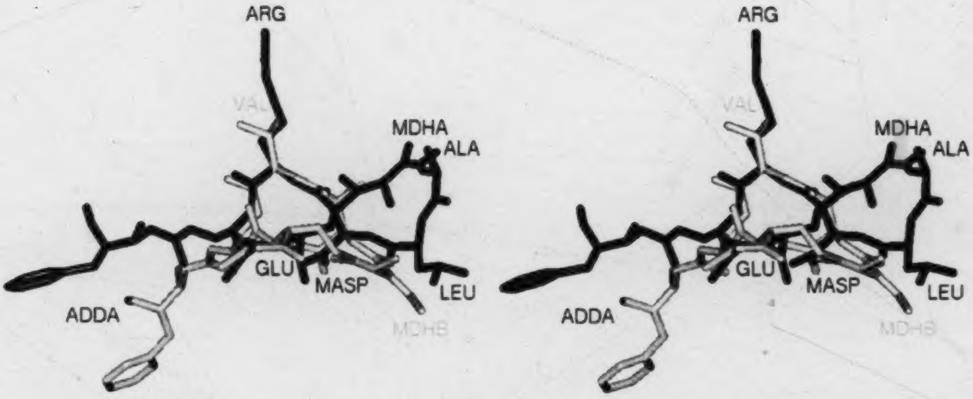


Figure 4- *Superimposition of microcystin-LR and motuporin average structures displaying only heavy atoms for clarity. Microcystin-LR is black and motuporin is grey. The black labels are for the residues in microcystin-LR or are shared with motuporin. The grey labels highlight the residues unique to motuporin. Superimposition of the Masp, arginine/valine, and Adda (except carboxyl carbon) backbone atoms was determined by these atoms having an unbreakable string of phi, psi, and omega angles of less than 40 degrees difference between microcystin-LR and motuporin.*

Comparison of Solution Structures to Modeled Structures

Our structure of microcystin-LR based on experimental NMR data was found to be different from the modelled structures published previously[14-16]. The main difference is that all modelled structures have a rectangular, planar cyclic ring while in our experimentally derived structure the ring forms a saddle shape. Also, the Adda side chain was found to be flexible and generally pointing away from the backbone whereas a previous model[14] pictured this long hydrophobic side chain coming back towards the backbone forming a U-shape with the phenyl group lying beneath the center of the ring.

While no previous modelled three dimensional structure of motuporin exists, a computer model has been proposed for nodularin[14]. Nodularin closely corresponds to motuporin with an Arg in place of the Val. In the nodularin model, the backbone is rectangular and planar. Our results show that motuporin has a non-planar ring. Further, it appears that the Arg to Val substitution from microcystin-LR to motuporin makes no difference to the shape of the backbone. It is thus implied that nodularin would not be planar but instead have a backbone shape like motuporin. Also, the adda side chain in nodularin would be flexible as in motuporin.

Comparison of the Solution Structures of Microcystin-LR and Motuporin

These toxins have a highly related primary sequence (Figure 1) with the main difference being that the ring of microcystin has two more amino acids than motuporin. In the NMR structures the backbone of the two peptides each resembled a saddle shape (Figures 3A and 3B). The Glu carboxyl, Masp carboxyl, and Adda side chain were found to be in a similar position relative to the ring in both microcystin-LR and motuporin (Figure 4). Previous studies have confirmed the essential nature of the Adda and Glu residues[5,9,17]. Namely, structurally modified 6(Z)-Adda microcystins showed 100 times weaker activity than the maternal 6(E)-Adda microcystins[5]. Also, methyl esterification of the Glu carboxyl side chain reduced the toxicity of microcystin-LR in a mouse bioassay[17]. It is thus reasonable to suggest that the similarities in backbone shape and positioning of the negatively charged side chains and long hydrophobic group are responsible for both peptides being able to recognize and inhibit PP-1c and PP-2Ac with high affinity and potency. Initial recognition of and attraction to the enzyme may take place with the Adda group. In each cyclic peptide the Adda side chain is behind the saddle (Figure 4) and with its flexible nature could easily adapt to fit the same hydrophobic patch on the phosphatase. Additionally, an ionic interaction between the peptide's negatively charged carboxyl groups of the Glu and/or Masp and a positively charged amino acid side chain of the phosphatase could contribute to the inhibitory binding. Since the positioning of either the Glu or Masp carboxyl groups is similar when comparing peptides, from a structural viewpoint either one or both could interact with the positive charge from the phosphatase. If both carboxyls were involved, the positive charge would probably lie underneath the ring.

Whilst microcystin-LR and motuporin exhibit many structural similarities, they differ in the relative positioning of the Mdha in microcystin-LR verses its counterpart residue Mdhb in motuporin (Figure 4). Mdha is known to be the residue involved in the covalent linkage of microcystin-LR to PP-1c (C.F.B.H., manuscript in preparation). The structural difference between these toxins which we have identified may explain why microcystin-LR is able to undergo a secondary time-dependent covalent interaction with PP-1c and PP-2Ac. The mechanism of covalent modification of PP-1c and PP-2Ac by microcystin-LR likely involves a Michael addition reaction between Mdha and a nucleophilic residue from the protein phosphatase (See Chapter 3 Page 52 Figure 5A for diagram of Cys 273-Mdha covalent link). While motuporin does possess a similar electrophilic center with the Mdhb residue, the presence of an additional methyl group attached to the beta carbon may conceivably sterically inhibit the covalent linkage of Mdhb to PP-1c. However, the difference in the relative position of Mdha verses Mdhb in the cyclic peptide backbones provides a more plausible reason for why microcystin-LR binds covalently to PP-1c and PP-2Ac whilst motuporin does not (Figure 4). Mdha is at the top of the front part of the saddle in microcystin-LR with the alpha carbon to beta carbon pointing upwards. The Mdhb in motuporin relative to Mdha is below and shifted into the position of Leu in microcystin-LR, suggesting that the significant relative displacement between Mdhb side chain and the protein phosphatases prevents the formation of a covalent link. Of course, it is possible that the structures of these toxins are different when bound to their target proteins as was observed for the cyclosporin/cyclophilin complex[18].

The inability of motuporin and nodularin to form a covalent link to PP-1c and PP-2Ac may ultimately explain why the nodularins may be able to function as carcinogens. Since these toxins form no covalent linkage to protein phosphatases, their Mdhb residues may be free to form direct adducts with DNA either when bound to protein phosphatase or following a dissociation event from these enzymes. This might explain why microcystin-LR appeared devoid of carcinogenic properties since it would be expected to remain covalently attached to PP-1c/PP-2Ac and therefore be unable to form an adduct with DNA via the Mdha residue.

References

1. Carmichael, W.W. The toxins of cyanobacteria. *Scientific American* **270(1)**, 78-86 (1994).

2. Honkanen, R.E. *et al.* Characterisation of microcystin-LR, a potent inhibitor of type-1 and -2A protein phosphatases. *J. Biol. Chem.* **265**, 19401-19404 (1990).

3. Yoshizawa, S. *et al.* Inhibition of protein phosphatases by microcystis and nodularin associated with hepatotoxicity. *J. Cancer Res. Clin. Oncol.* **116**, 609-614 (1990).

4. MacKintosh, C., Beattie, K.A., Klumpp, S., Cohen, P. & Codd, G.A. Cyanobacterial microcystin-LR is a potent and specific inhibitor of protein phosphatases 1 and 2A from both mammals and higher plants. *FEBS Lett.* **264**, 187-192 (1990).

5. Nishiwaki-Matsushima, R. *et al.* Structure-function relationships of microcystins, liver tumour promoters, in interaction with protein phosphatase. *Jpn. J. Cancer Res.* **82**, 993-996 (1991).

6. Holmes, C.F.B. & Boland, M.P. Inhibitors of protein phosphatase-1 and -2A; two of the major serine/threonine protein phosphatases involved in cellular regulation. *Current Opinion in Structural Biology* **3**, 934-943 (1994).

7. Eriksson, J.E. *et al.* Hepatocyte deformation induced by cyanobacterial toxins reflects inhibition of protein phosphatases. *Biochem. Biophys. Res. Commun.* **173**, 1347-1353 (1990).

8. Ohta, T. *et al.* Hyperphosphorylation of cytokeratins 8 and 18 by microcystin-LR, a new liver tumour promoter, in primary cultured rat hepatocytes. *Carcinogenesis* **13**, 2443-2447 (1992).

9. Rhinehart, K.L *et al.* Nodularin, microcystin and the configuration of adda. *J. Am. Chem. Soc.* **110**, 8557-8558 (1988).

10. Craig, M. *et al.* Identification and characterization of hydrophobic microcystins in canadian freshwater cyanobacteria. *Toxicon* **31(12)**, 1541-1549 (1993).

11. Ohta, T. *et al.* Significance of the cyanobacterial cyclic peptide toxins, the microcystins and nodularin in liver cancer. *Mutation Research* **292**, 286-287 (1993).

12. DeSilva, S.D. *et al.* Motuporin, a potent new protein phosphatase inhibitor from the Papua New Guinea sponge *Theonella swinhoei* Gray. *Tetrahedron Lett.* **33**, 1561-1564 (1992).

13. MacKintosh, C. *et al.* Inhibition of protein phosphatases by toxins: implications for health and an extremely sensitive and rapid method for toxin detection. *Proceedings of the Royal Society of Chemistry (U.K.)* **In Press**, 1994.

14. Lanaras, T., Cook, C.M., Eriksson, J.E., Meriluoto, J.A.O. & Hotokka, M. Computer modelling of the 3-dimensional structures of the cyanobacterial hepatotoxins microcystin-LR and nodularin. *Toxicon* **29(7)**, 901-906 (1991).

15. Quinn, R.J., Taylor, C., Suganuma, M. & Fujiki, H. The conserved acid binding domain model of inhibitors of protein phosphatases 1 and 2A: molecular modelling aspects. *Bioorganic & Medicinal Chemistry Letters* **3(6)**, 1029-1034 (1993).

16. Fujiki, H. & Suganuma, M. Tumor promotion by inhibitors of protein phosphatases-1 and -2A: the okadaic acid class of compounds. *Advances in Cancer Research* **61**, 143-194 (1993).

17. Stotts, R.R. *et al.* Structural modifications imparting reduced toxicity in microcystins from *Microcystis sp. Toxicon* **31**, 783-789 (1993).

18. Fesik, S.W. *et al.* NMR studies of [U-^{13}C]cyclosporin A bound to cyclophilin: bound conformation and portions of cyclosporin involved in binding. *Biochemistry* **30**, 6574-6583 (1991).

Thesis Chapter 3

A Molecular Basis for Different Interactions of Marine Toxins With Protein Phosphatase-1: Molecular Models for Bound Motuporin (Nodularin-V), Okadaic Acid, Calyculin A, and the Microcystins

This chapter published as Bagu J.R., Sykes B.D., Craig M.M., Holmes C.F.B. (1997) *J Biol Chem* **272**, 5087-5097.

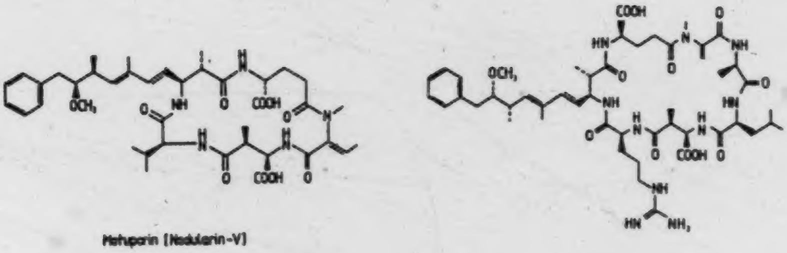
Summary

The hepatotoxic cyclic heptapeptide microcystins are powerful liver tumor promoters. Biochemically they act as potent inhibitors of protein phosphatase-1 and -2A, and are able to covalently bind to these enzymes. The hepatotoxic cyclic pentapeptide nodularins are also liver tumor promoters and potent inhibitors of protein phosphatase-1 and -2A, but they do not covalently interact with the protein phosphatases. The nodularins may additionally possess carcinogenic properties. Okadaic acid and calyculin A are other inhibitors of protein phosphatase-1 and -2A. The conformation of microcystin-LR has been determined in solution[1] and bound to the catalytic subunit of protein phosphatase-1[2]. The free NMR solution structure of microcystin-LR is remarkably similar to the bound crystal structure of this cyanobacterial toxin. We show herein that the NMR solution structure of a novel hydrophobic microcystin congener (microcystin-LL) is also very similar to the solution structure microcystin-LR. Using Metropolis Monte Carlo modelling to dock the solution structure of microcystin-LL onto the crystal structure of protein phosphatase-1, we found relatively minor differences in the structural orientation of this toxin when compared to bound microcystin-LR in the crystal structure complex. We have further exploited this finding by docking the hydrophobic toxin motuporin (nodularin-V) onto protein phosphatase-1, showing that it also occupies a similar position to microcystin-LR when bound to protein phosphatase-1. However, there is a striking difference in the position of the N-methyldehydroalanine residue in microcystin-LR relative to the N-methyldehydrobutyrine residue in motuporin (nodularin-V). This difference in orientation provides a molecular explanation for why the nodularins may be incapable of forming a covalent linkage with protein phosphatase-1. Furthermore, the position of the N-methyldehydrobutyrine residue in motuporin is at the surface of the protein phosphatase-toxin complex which may facilitate chemical interaction with a further distinct macromolecule(s) possibly relating to its carcinogenic properties. We have also used the previously published free crystal structures of okadaic acid[3] and calyculin A[4] in docking them to the crystal structure of protein phosphatase-1[2]. Both of these inhibitors are similar to microcystins and motuporin in their tertiary structure and binding position relative to protein phosphatase-1. The bound model for okadaic acid is also useful in explaining the previously published work in determining what regions of okadaic acid/protein phosphatase-1 are important (or not) for their interaction.

Introduction

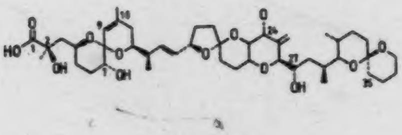
The microcystin and nodularin classes of peptide hepatotoxins are metabolites of cyanobacteria in the genera *Microcystis*, *Anabaena*, and *Nodularia* that grow worldwide in fresh and brackish water[5]. These toxins are responsible for extensive wildlife fatalities, and adverse effects on human health have also been recognized in countries where drinking water supplies contain cyanobacteria. It has been shown that microcystin-LR is a potent inhibitor of the catalytic subunits of protein phosphatase-1 and -2A (PP-1c and PP-2Ac) as well as a powerful tumor promoter[6-10]. Inhibition of these enzymes in the liver is apparently associated with hepatocyte deformation due to reorganization of microfilaments[11]. Liver tumor promotion may be linked to the ability of this cyclic peptide to promote hyperphosphorylation of cytokeratins associated with morphological changes in rat hepatocytes[12].

A unique chemical feature that characterizes the microcystin/nodularin peptide hepatotoxins is the presence of the unusual C_{20} β-amino acid [2S, 3S, 8S, 9S]-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid (Adda)[13]. To date over 40 different microcystins have be characterized[14]. Within the microcystin family, members differ from each other in the nature of two variable L-amino acids indicated by suffix letters (eg. L=Leu, R=Arg), and in the presence or absence of methyl groups on the D-erythro-β-methylaspartic acid (Masp) and/or N-methyldehydroalanine (Mdha) residues (Fig. 1). Despite these differences no loss of protein phosphatase inhibition or tumor promotion has been observed. Nodularins are structurally related cyclic pentapeptides (Fig. 1) that inhibit PP-1c and PP-2Ac with similar potency to microcystins and are also powerful tumor promoters[15]. The relatively hydrophobic cyclic pentapeptide motuporin (termed nodularin-V in Fig. 1), which was isolated from the marine sponge *T. swinhoei*, is the only nodularin analog known[16].

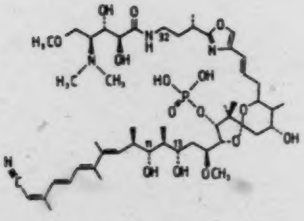


Motuporin [Nodularin-V]

Microcystin-LR



Okadaic acid



Calyculin A

Figure 1- *Primary sequences of microcystin, nodularin, okadaic acid, and calyculin A. Residues for microcystin-LR (top,right) are β-linked D-erythro-β-methylaspartic acid (Masp); L-Arginine; β-[2S, 3S, 8S, 9S]-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid (Adda); γ-linked D-glutamic acid (Glu); N-methyldehydroalanine (Mdha); D-alanine; and L-Leucine. Motuporin aka. nodularin-V (top,left) has only 5 residues which are Masp; L-Valine; Adda; Glu; and N-methyldehydrobutyrine (Mdhb).*

Although sharing similar biological properties, important functional differences between the microcystins and nodularins have been identified. One difference is in the interaction with PP-1c and PP-2Ac. Although both toxins initially bind non-covalently and inhibit these enzymes, microcystin-LR, -LA, -LL undergo a secondary time-dependent interaction with the phosphatase[17]. This secondary interaction results in a covalent linkage causing irreversible modification of PP-1c/PP-2Ac. In contrast, nodularin or motuporin do not bind covalently to PP-1c/PP-2Ac after inhibiting it. A second difference is that in addition to acting as a tumor promoter, nodularin may also act as a carcinogen/tumor initiator[15].

Okadaic acid and its analogs are polyether-like compounds (Fig. 1), and as with microcystins and nodularins are potent inhibitors of PP-1c/PP-2Ac and powerful tumor promoters[10]. These toxins produced by the unicellular dinoflagellates *Dinophysis* and *Prorocentrum* are partly responsible for diarrhetic shellfish poisoning. Okadaic acid has also been used extensively as a research tool in studying reversible protein phosphorylation in many cellular processes[18]. Calyculin A (Fig. 1) and its analogs isolated from the marine sponge *Discodermia calyx* are effective inhibitors of PP-1c and PP-2Ac, being tumor promoters as well[19]. Interestingly, the chemical stuctures of the microcystin/nodularin, okadaic acid, and calyculin A differ from each other (Fig. 1) despite having the similar property of inhibiting PP-1c/PP-2Ac related to their tumor promotion activity.

Previously, the solutions structures of microcystin-LR and motuporin (nodularin-V) were compared with each other so as to give insights into how the residues and the tertiary structure were important for binding to PP-1c/PP-2Ac[1]. Comparison of these solution structures which revealed how the Mdhb residue in motuporin was relatively significantly displaced from the Mdha residue in microcystin-LR, thus providing a preliminary explanation for why microcystins are capable of covalently binding to the protein phosphatases while motuporin lacks the ability[1].

Elucidation of the crystal structure of microcystin-LR bound covalently to PP-1c[2] provided a detailed view of the key interactions between microcystin-LR and PP-1c. The elucidation of the bound form of microcystin-LR in complex with PP-1c[2] allows for a comparison of the structures of unbound and bound microcystin-LR. Accordingly, I show that the published free solution NMR structure of microcystin-LR does not change significantly in conformation upon inhibiting and binding covalently to PP-1c. This may be a factor contributing to the high affinity of microcystin-LR for PP-1c. and other inhibitors of PP-1c to be docked with the crystal structure of PP-1c.

Given the remarkable similarity between the free solution structure of microcystin-LR[1] and the bound crystal structure[2], we predict that the other microcystins will have similar free and bound structures. In accordance with this hypothesis, we determined the solution structure of the hydrophobic microcystin congener microcystin-LL and docked this structure (using Metropolis Monte Carlo modelling) onto the crystal structure of PP-1c. Further, because nodularins are structurally related to microcystins[1], we also assume that the nodularins will largely retain their solution structure upon binding to PP-1c, and have also docked motuporin onto the protein phosphatase. A comparison of the complexes of microcystin-LR, microcystin-LL, and motuporin with PP-1c sheds light upon the different molecular mechanisms of their interaction with the phosphatase. These comparisons are used to explain why the microcystins covalently react and the nodularins do not, and provide a mechanistically appealing hypothesis for explaining the carcinogenic properties of the nodularins. We have also taken previously published free crystal structures of okadaic acid[3] and calyculin A[4] and docked them to PP-1c in the same manner as microcystin-LL and motuporin. These docked models of okadaic acid and calyculin A in comparison to microcystins and motuporin account for how these diverse toxins are able to inhibit PP-1c.

Methods

Microcystin-LL was purified from algal extracts as previously described[14,16]. The sample used for NMR was 1mM dissolved in a 10mM potassium phosphate, 50mM sodium chloride buffer with 80-90% H2O/10-20% D2O at pH 7. 2,2-dimethyl-2-silapentane-5-sulfonic acid (0.1mM) was added as a ^{1}H NMR chemical shift standard.

^{1}H NMR spectra were recorded at 500MHz on a Varian VXR-500 NMR spectrometer. Proton NMR resonance assignments were made using standard sequential assignment methods with double quantum filtered COSY, TOCSY, and NOESY two dimensional ^{1}H NMR spectra and were similar to microcystin-LR and motuporin assignments[1]. ^{1}H-^{1}H internuclear distance restraints were obtained from two dimensional ^{1}H NOESY taken with a mixing time of 150ms so as to minimize spin diffusion. $^3J_{NH\alpha CH}$ and $^3J_{CH\text{-}CH}$ coupling constants were obtained by curvefitting one dimensional ^{1}H NMR spectra. Assignments were made at 25°C and 5°C, with structural information recorded at 5°C. Calibration of the NOEs was accomplished using the known distance of 1.8Å between the HB protons of the sp^2 carbon of the Mdha side chain to scale the experimental crosspeak intensity. Sixty structures were generated using the distance geometry program DGII in the program InsightII version 2.3 (Biosym Technologies Inc.) with 58 of the 60 structures converging into one fold and the other 2 discarded due to violations. For this cyclic peptide the vicinal angle restraints were very important in restraining the structure of the ring. Distance and angle violations were analyzed using the NMR refine module in InsightII. There were no consistently violated distance restraints over 0.1Å and angles fell within of the defined restraint range in almost all cases for the 58 converging structures. An average structure was then generated from the 58 structures, and was subjected to constrained minimization to correct bond distances that had been distorted through averaging.

Metropolis Monte Carlo docking of microcystin-LL and motuporin was accomplished using the Monte Carlo macro in the program Insight II version 2.3 (Biosym Technologies Inc.). For microcystin-LL, the minimized average structure was used for docking with the PP-1c crystal structure[2]. The PP-1c crystal structure was modified to facilitate docking by removing the bound microcystin-LR and adding protons. The starting position for microcystin-LL relative to PP-1c was obtained by superimposing microcystin-LL onto the bound X-ray crystal structure of microcystin-LR. This was accomplished by superimposing the identical backbone atoms of the "docking" phosphatase structure onto the original phosphatase structure that still had microcystin-LR attached, then superimposing microcystin-LL onto the crystal microcystin-LR, then removing the original phosphatase+microcystin-LR leaving docking phosphatase with microcystin-LL in its starting position. From this single starting point, the Monte Carlo macro then preformed docking calculations for 2000 trials at a temperature of 50^{0}K. Other temperatures (1^{o}K and 100^{o}K) were tried but varying temperature did not make any difference in the positioning of the final docked structure of microcystin-LL if the number of trials was equivalent between them. Furthermore, different numbers of trials were used (100, 1000, and 2000) at 50°K with all three amounts of trials resulting in a similarly docked microcystin-LL. Because 2000 trials resulted in the lowest energy for the final structure this was chosen for the best number of trials. For motuporin, again the minimized average solution structure was used for docking along with the PP-1c with protons but lacking the bound microcystin-LR. Superimposition of the motuporin onto the bound microcystin-LR was accomplished in the same manner as microcystin-LL. Using what was learned with respect to temperature and number of trials in the case of microcystin-LL, docking of motuporin using the Monte Carlo macro was at 50^{0}K with 2000 trials.

To carry out molecular modeling studies with okadaic acid[3] and calyculin A[4] the free crystal structures were retrieved from the Cambridge Structural Database (*Cambridge Crystallographic Data Center*)[20]. In the case of okadaic acid, the bromobenzyl region of the crystal structure was removed so as to use native okadaic acid in the docking process. Both okadaic acid and calyculin A were docked in the same manner as microcystin-LL and motuporin. The starting point used for okadaic acid and calyculin A was the visual superimposition of the individual toxin onto bound microcystin-LR.

As a rigorous test of our procedues, the minimized average solution structure of microcystin-LR was docked onto the crystal structure of PP-1c in the same manner and starting point as microcystin-LL and motuporin. The docked structure of microcystin-LR hardly moved from its starting position, indicating that the Metropolis Monte Carlo macro was successful in keeping the free solution structure of microcystin-LR at the same binding position observed in the bound X-ray crystallography form of microcystin-LR[2] (data not shown). In addition, we calculated the root mean square deviation (RMSD) change in atomic positions from the original manually positioned toxins to the final model after docking. These data (Table 1) show that the RMSD for all atoms between starting and docked structures for each toxin varies between 0.68 Å for microcystin-LR and 2.36 Å for okadaic acid.

TABLE 1- RMSD* between starting and final docked structures for each toxin

Toxin	RMSD (Å)
Microcystin-LR	0.683
Microcystin-LL	1.56
Motuporin	1.96
Okadaic Acid	2.36
Calyculin A	1.26

*Root-mean-square deviation for all atoms.

As a further test of the robustness of our docking methods, we employed our Metropolis Monte Carlo procedures for the converging 58 NMR structures of microcystin-LL. Although most of these structures do indeed dock in a manner similar to that of the average structure, to be objective we focused particularly on the six most disparate microcystin-LL NMR structures. The disparate nature of these distorted structures was predicated largely by the flexibility of their Adda side chains in solution. Therefore, even though the main ring of these toxins might have to move an unrealistic amount to compensate for an unusually positioned Adda side chain, these six most distorted toxin structures still fitted well, wherein the backbone RMSD (comparing docked NMR calculated structures with the docked average solution structure for microcystin-LL) averaged 1.28 Å (+/- 0.54 S.D.).

Results

Comparison of free NMR microcystin-LR with bound crystal microcystin-LR

We have compared (Fig. 2) the minimized average free solution structure of microcystin-LR (as determined by NMR spectroscopy[1]) and the structure of microcystin-LR bound to the phosphatase PP-1c (as determined by X-ray crystallography[2]). The root mean square deviation (RMSD) of the backbone atoms superimposed on each other is 0.65Å (the backbone atoms in all cases include the additional backbone carbon atoms in the Masp, Adda, and Glu residues). The two structures are similar indicating that microcystin-LR retains its free conformation while inhibiting and covalently binding to PP-1c. There are relatively minor differences in the orientation of the arginine and Mdha side chains between the two structures. The minor difference in the location of the arginine side chain may be accounted for by the flexibility of the arginine side chain in solution and by the fact that the arginine side chain in the crystal structure was disordered so that its position could not accurately be determined. In addition, the change in position of the Mdha side chain at the front of the saddle (where it points upwards in the free solution structure and downwards in the bound crystal structure, as depicted in Fig. 2) is explained by the covalent binding that occurs between the Mdha residue and Cys 273 of PP-1c. The Mdha side chain must turn downwards when covalently linking to the enzyme. However, the Adda side chain seems to be in a similar position when comparing both structures. In solution, the Adda side chain is flexible[1] with the position used in Fig. 2 being that of the average structure. In the crystal structure[2] the Adda side chain has lost some of its flexibility and is clearly seen, unlike the arginine side chain.

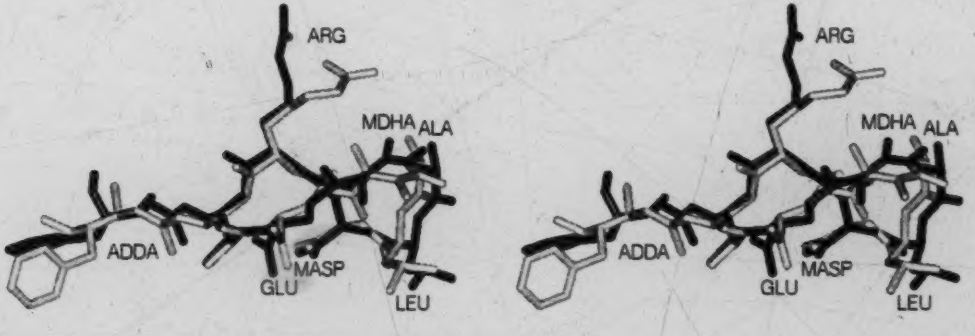


Figure 2- *Stereoview of comparison between the free solution microcystin-LR and bound crystal microcystin-LR. The free average minimized NMR structure of microcystin-LR is black, while the bound crystal structure is in grey. For clarity, neither structure has hydrogen atoms. The RMSD of the backbones superimposed on each other is 0.65Å.*

The family of solution structures of microcystin-LL determined by ^{1}H NMR is shown in Fig. 3A. The RMSD of the backbone atoms relative to the average structure is 0.57Å +/- 0.10Å. In the free solution structure of microcystin-LL, the cyclic backbone is a relatively rigid saddle and the Adda side chain is highly flexible. The positioning of the Glu and Masp carboxyl groups are in such a way as to have their negative charges situated below the saddle backbone. The Mdha residue, which covalently links to PP-1c, is located at the top, front part of the saddle from the perspective in Fig. 3A. With these features, the minimized average structure of microcystin-LL is highly comparable to the minimized average solution structure of microcystin-LR[1] (Fig. 3B), with the RMSD of the two backbones relative to each other being 0.77Å. Clearly, the substitution of the positively charged Arg with the hydrophobic Leu results in no overall three dimensional structural changes. Given that microcystin-LR retains its solution structure upon binding to PP-1c, we postulate that the similar microcystin-LL would do likewise. On this basis, the minimized average free solution structure of microcystin-LL was docked onto the crystal structure of PP-1c[2] using Metropolis Monte Carlo. Microcystin-LL retains a highly similar position as the bound crystal structure of microcystin-LR relative to PP-1c. This is illustrated by the presence of virtually identical protein phosphatase residues in close proximity (within 4Å) to microcystin-LL as to bound microcystin-LR (see Fig. 5A).

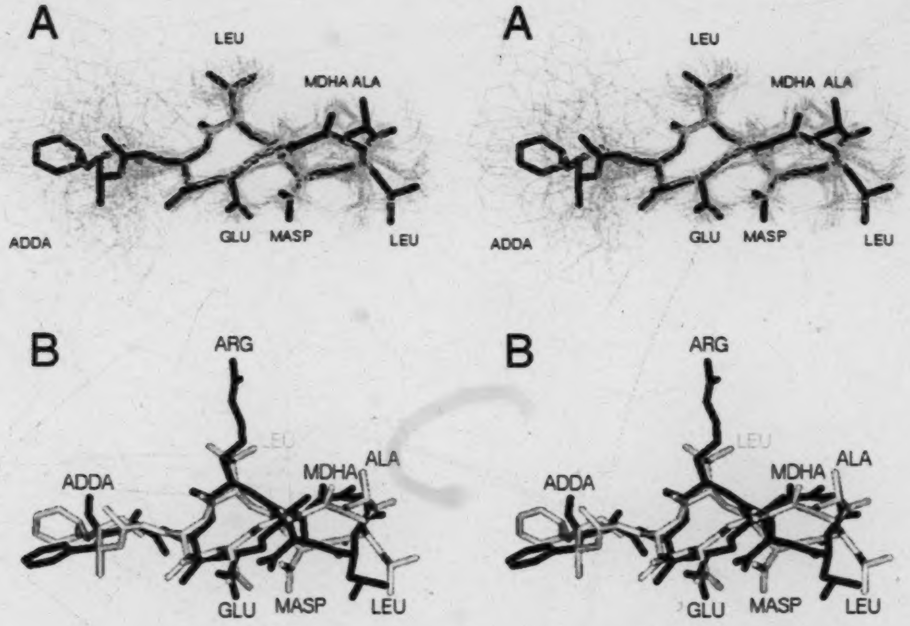


Figure 3- *Solution structure of microcystin-LL in stereo. Ensemble of calculated NMR solution structures of microcystin-LL. The 58 calculated DGII structures are shown in grey; the average minimized structure is in black. The backbone RMSD is 0.57Å +/- 0.10Å of the calculated structures relative to the average structure. Hydrogen atoms are not shown.* ***B.*** *Superimposition of free solution microcystin-LR and free solution microcystin-LL in stereo. The average minimized free NMR structure of microcystin-LR is shown in black, with the grey structure being the averaged minimized free NMR structure of microcystin-LL. The difference in sequence is that microcystin-LR has an arginine (label in black) where microcystin-LL has a leucine (label in grey). The backbone RMSD with respect to the two structures is 0.77Å. No hydrogen atoms are shown.*

Superimposition of free NMR motuporin onto crystal PP-1c

The average free minimized solution structure of motuporin[1] is superimposed onto the bound crystal structure of microcystin-LR[2] in Fig. 4. The backbone segment of the Masp-Val-Adda atoms of motuporin (excluding the Adda carbonyl carbon) was superimposed onto the corresponding Masp-Arg-Adda atoms of bound crystal microcystin-LR. This superimposition was chosen because this range of atoms had similar phi, psi, and omega angles and is identical to the one comparing the average free solution structures of microcystin-LR and motuporin previously published[1]. The RMSD of these superimposed atoms between the two structures is 0.46Å. These data reaffirm the similarities in the backbone atoms of residues Masp, arginine/valine, and Adda, the positioning of the carboxyls (Masp and Glu), and the positioning of the Adda side chain between microcystin-LR and motuporin. Once again, the difference in the Mdha/Mdhb section of the inhibitors results in a major displacement of the Mdhb residue in motuporin relative to the Mdha residue in microcystin-LR1 shown here to be 7.13Å with respect to the beta carbons of each residue.

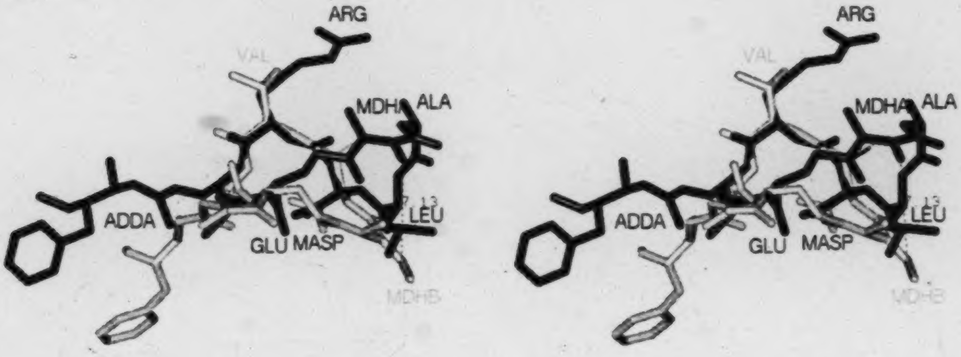


Figure 4- Superimposition of the free average minimized solution structure of motuporin onto the bound crystal structure of microcystin-LR in stereo. Motuporin and its labelled residues belonging to it are in grey, the black structure and black labelled residues belong to microcystin-LR, and the some of the black labelled residues are common between the two structures. The distance of 7.13Å is between the beta carbons of the Mdha residue in microcystin-LR and the Mdhb residue in motuporin. No hydrogens atoms are shown.

Assuming that motuporin (like microcystin-LR) retains its free solution structure upon binding to PP-1c, we docked the free averaged minimized solution structure of motuporin[1] onto the crystal structure of PP-1c[2] by Metropolis Monte Carlo (Fig. 5B). The distance between the beta carbon of the Mdhb residue to the sulphur atom on the Cys 273 side chain is 10Å. Similar to the results achieved with microcystin-LL[1] docking, motuporin differs only slightly in position relative to bound microcystin-LR. This model predicts similar protein phosphatase residues to be within 4Å of the toxin when compared to bound microcystin-LR (Fig. 5A) and docked microcystin-LL, but lacks some proximity connections (not within 4Å of) particularly with the PP-1c residues 274-276.

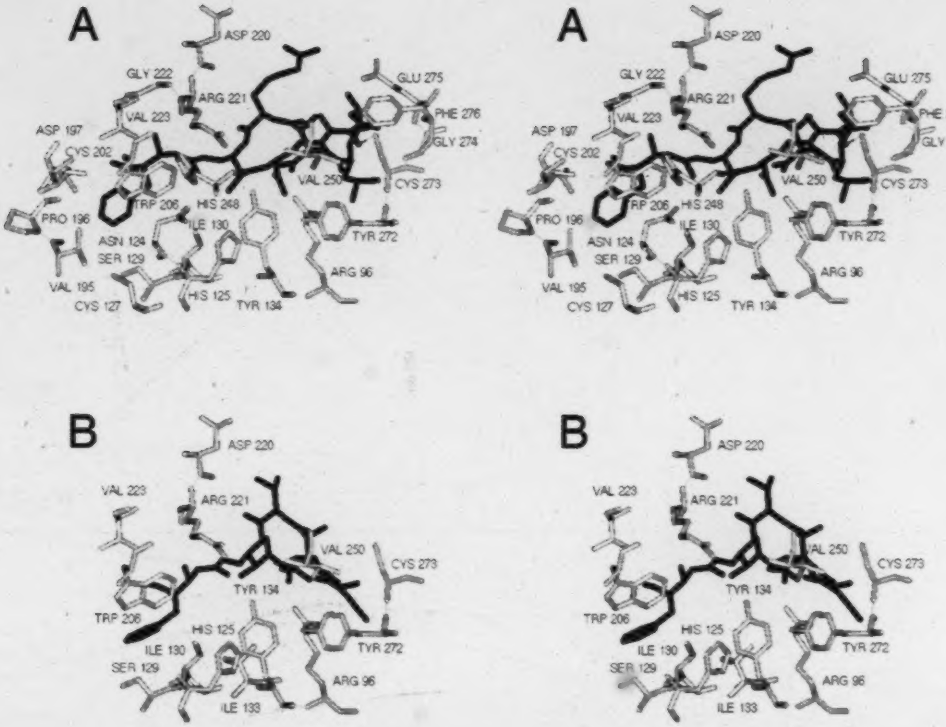


Figure 5A- *Bound crystal microcystin-LR complexed with protein phosphatase-1c in stereo (top). Bound microcystin-LR is in black with the phosphatase residues in grey.* ***B****. Metropolis Monte Carlo docking of the free average minimized structure of motuporin onto the crystal structure of PP-1c in stereo (bottom). Motuporin is in black with PP-1c residues in grey. Only phosphatase residues that have atoms within 4Å of either microcystin-LR or motuporin are displayed. The same orientation is shown for the phosphatase residues in both cases. As well, A and B have no hydrogen atoms present.*

An overall view of the crystal structure of microcystin-LR bound to PP-1c[2] (Fig. 6A) is compared to the model for motuporin docked onto the crystal structure of PP-1c (Fig. 6B) with the phosphatases being oriented identically. A remarkable difference between the two phosphatase-toxin complexes is that the microcystin Mdha residue which is covalently bonded to the Cys 273 is relatively buried compared to the equivalent Mdhb residue in motuporin. The general lack of phosphatase residues surrounding the Mdhb residue in motuporin indicates that this residue is highly surface accessible and in fact protrudes from the phosphatase-toxin complex.

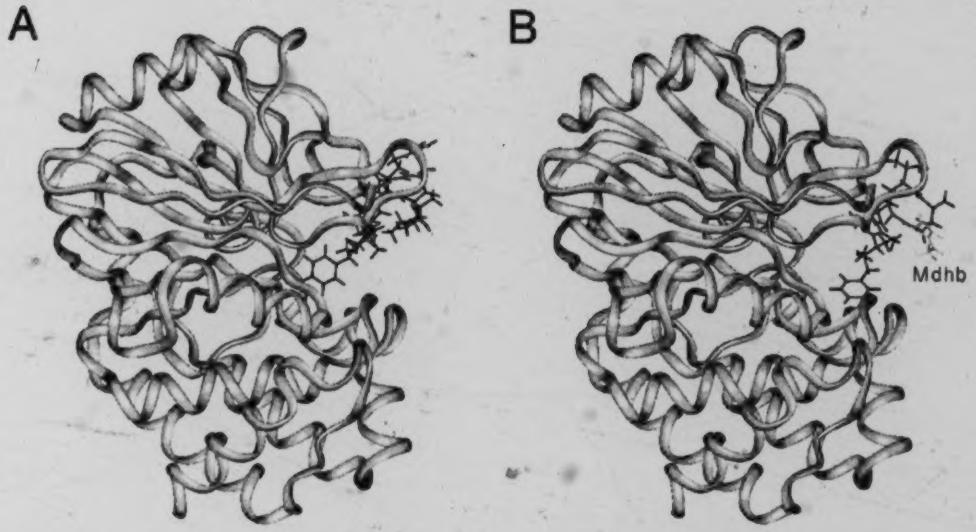


*Figure 6A- Overall view of microcystin-LR and protein phosphatase-1c crystal structure complex (left). PP-1c backbone is in a ribbon and bound microcystin-LR is black. **B.** Same view as **A.** except have docked motuporin instead of bound microcystin-LR (right). Free average minimized solution structure of motuporin is in black except the Mdhb residue which is in grey and also labelled. PP-1c backbone is in a ribbon.*

Superimposition of free crystal structures of OA and calyculin A onto bound PP-1c

Fig. 7A depicts the docked free crystal structure of okadaic acid compared with the bound crystal structure of microcystin-LR. The orientation shown is the positions of the toxins relative to each other when bound to PP-1c. Interestingly, there are similarities in the three dimensional structure and positioning of docked okadaic acid compared to bound microcystin-LR. The hydrophobic end of okadaic acid overlaps with the hydrophobic Adda residue of microcystin-LR. The tertiary structure of okadaic acid continues to follow microcystin-LR up to the arginine residue, but then extends over the top of the microcystin saddle to once again overlap with microcystin at the Mdha residue. Finally, the C1 carboxyl end finishes up in a position very near the Glu residue of microcystin-LR. Thus, the hydrophobic sections and carboxyl groups of each toxin overlap, with there being a resemblance in backbone shape. Fig. 7B is a view of docked okadaic acid with the surrounding PP-1c residues that are within 4Å of it. The hydrophobic section of okadaic acid is able to fit into the hydrophobic pocket similar to the Adda residue of microcystin and motuporin. The other carboxyl end of okadaic acid is near to Arg 96 and Tyr 272 of PP-1c, like the Glu and Masp carboxyl groups of microcystin and motuporin. The data in Fig.7 underscore a predicted structural similarity among OA, microcystin-LR, microcystin-LL, and motuporin such that these toxins all appear to have many PP-1c contact residues in common. Therefore, okadaic acid appears to not only bind to the same region as microcystin and motuporin, but may undergo similar interactions with the phosphatase residues present in the microcystin binding pocket.

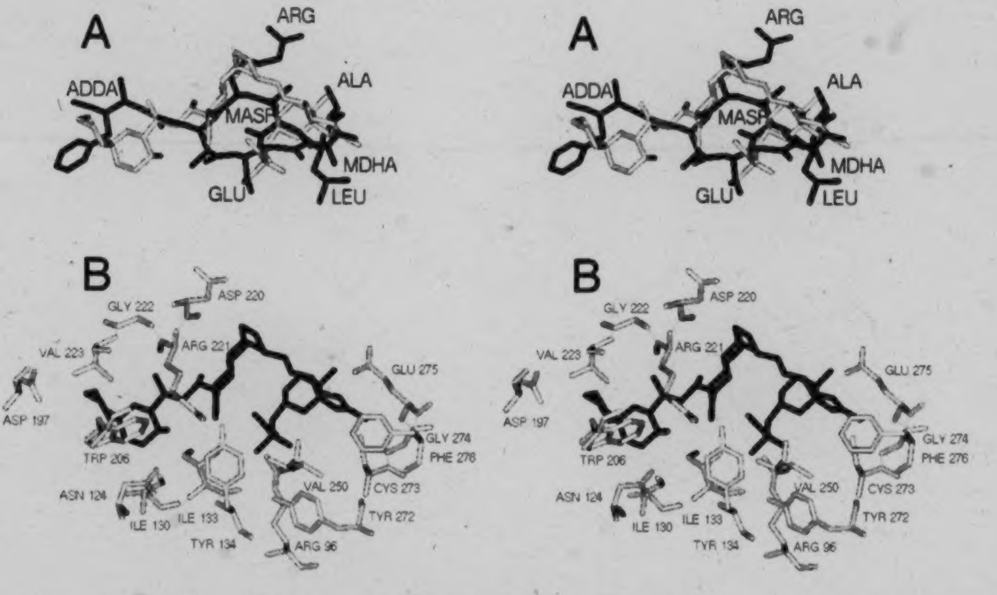


Figure 7A- *View of docked okadaic acid (grey) with the bound crystal structure of microcystin-LR (black). These two toxins are positioned to show how they are in comparison to each other when bound to PP-1c. The superimposed phosphatases are not shown.* ***B.*** *Same position of okadaic acid as* ***A.*** *except now showing the surrounding PP-1c residues within 4Å. PP-1c residues are shown in grey while okadaic acid is in black. Fig. 7A and 7B are both in stereo and neither show hydrogens for clarity.*

The docked free crystal structure of calyculin A shown in Fig. 8A with the bound crystal structure of microcystin-LR is the exact oriention relative to PP-1c as bound microcystin-LR. Like okadaic acid, the chemically dissimilar calyculin A resembles bound microcystin-LR in terms of their three dimensional structure and position in relation to PP-1c. Starting at the hydrophobic end of calyculin A which matches the Adda residue in microcystin-LR the two structures remain in agreement until calyculin A reaches the Adda backbone atoms, where it then goes underneath the saddle skipping arginine until it reaches the Masp residue. From there, it follows the saddle backbone of microcystin-LR going from Masp to leucine to alanine all the way back to the arginine side chain. The negative charged phosphate group of calyculin A is located above the Masp group of microcystin-LR but they still somewhat close to each other. Fig. 8B shows the surrounding PP-1c residues within 4Å of calyculin A. Calyculin A is close to many of the same PP-1c residues as the other inhibitors, indicating that it binds in the same pocket. It also has a hydrophobic and negatively charged regions that interact with the hydrophobic and basic regions of PP-1c in the same manner as the other inhibitors. It is predicted that calyculin A would bind in the same pocked as the other PP-1c inhibitors.

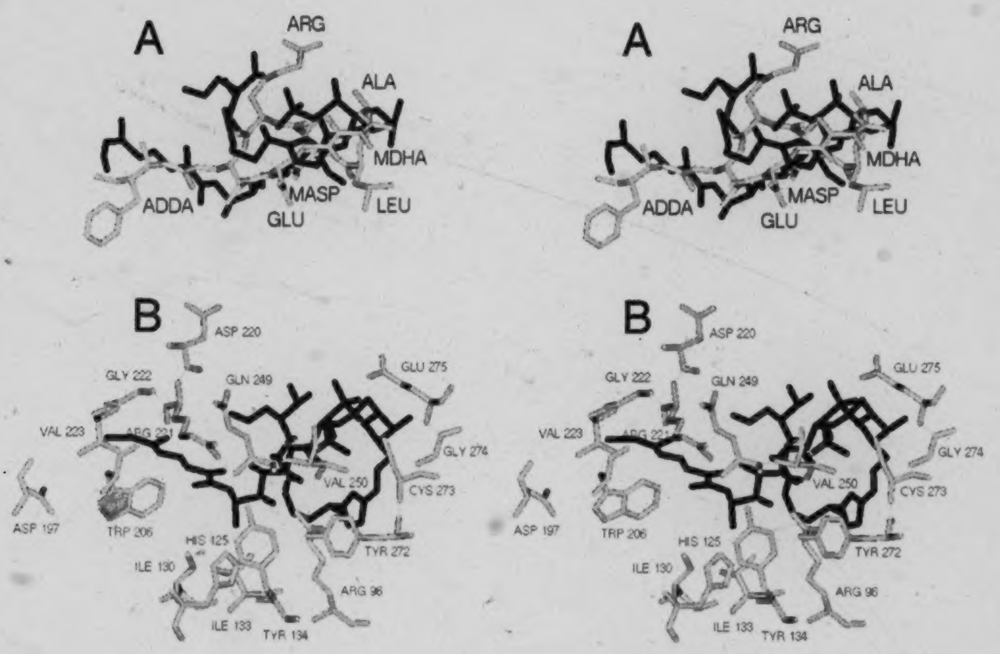


Figure 8A- *View of docked calyculin A (grey) with the bound crystal structure of microcystin-LR (black). The two toxins are overlapped in the same relative position to PP-1c, with the superimposed phosphatases not shown.* ***B.*** *Same position for calyculin A as* ***A.*** *except now showing the surrounding PP-1c residues within 4Å. Calyculin A is in black, while PP-1c residues are in grey. Fig. 8A and 8B are both in stereo.*

Determination of predicted molecular surfaces of free and bound marine toxins

The nature of the molecular surfaces of microcystin-LR, motuporin, OA, and calyculin A in both their unbound (Fig. 9) and PP-1c bound/docked (Fig. 10) conformations was calculated using the GRASP program. The surface of each toxin was examined according to its electrostatic surface potential (in the bound form) and its distance from the surface of PP-1c (in the bound/docked form). These experiments revealed a strong similarity in the tertiary structures of the toxins and prompted a further GRASP analysis of the molecular surfaces of PP-1c involved in toxin binding (Fig. 11). Interestingly, the predicted molecular surfaces of PP-1c closely involved in binding microcystin-LR (Fig. 11A) and OA (Fig. 11C) were the most strikingly similar. These include the region on PP-1c (represented in white, 0-1.5 Å distance) comprising Tyr-272 and Cys-273. Notably, this region (represented in red in Fig. 11B) is not predicted to be close to Mdhb in motuporin and appears to be less important in interacting with this toxin. These data are again consistent with the idea that motuporin does not interact covalently with PP-1c.

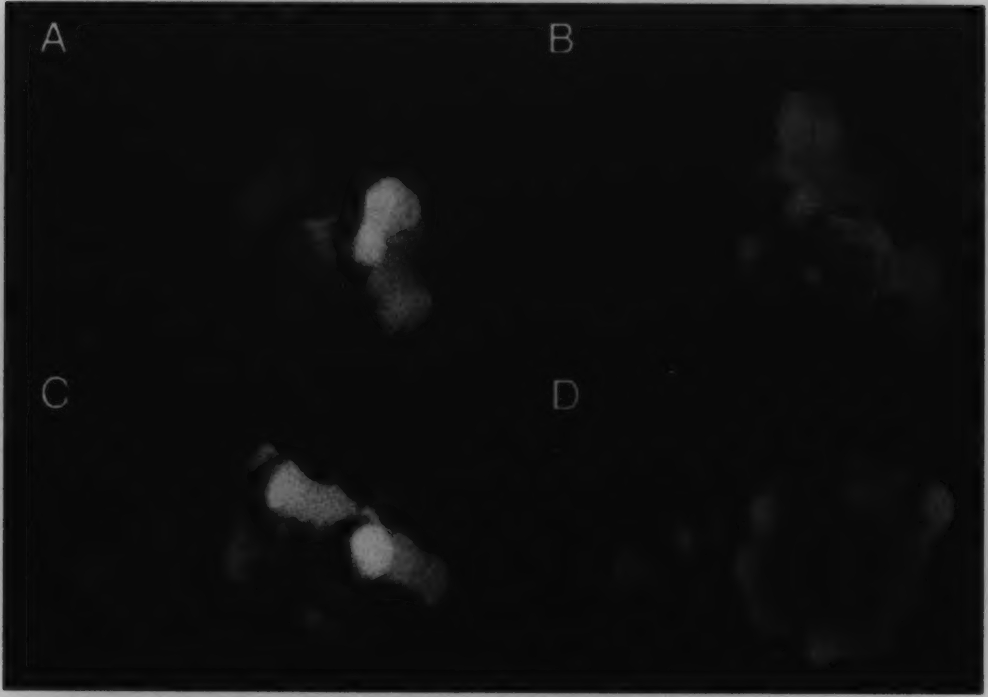


Figure 9- Molecular surface of free A. average solution structure of microcystin-LR, B. average solution structure of motuporin, C. crystal structure of okadaic acid, and D. crystal structure of calyculin A calculated by GRASP (Ref. 33). The surface is colored according to electrostatic surface potential where blue is positive, white is neutral, and red is negative potential.

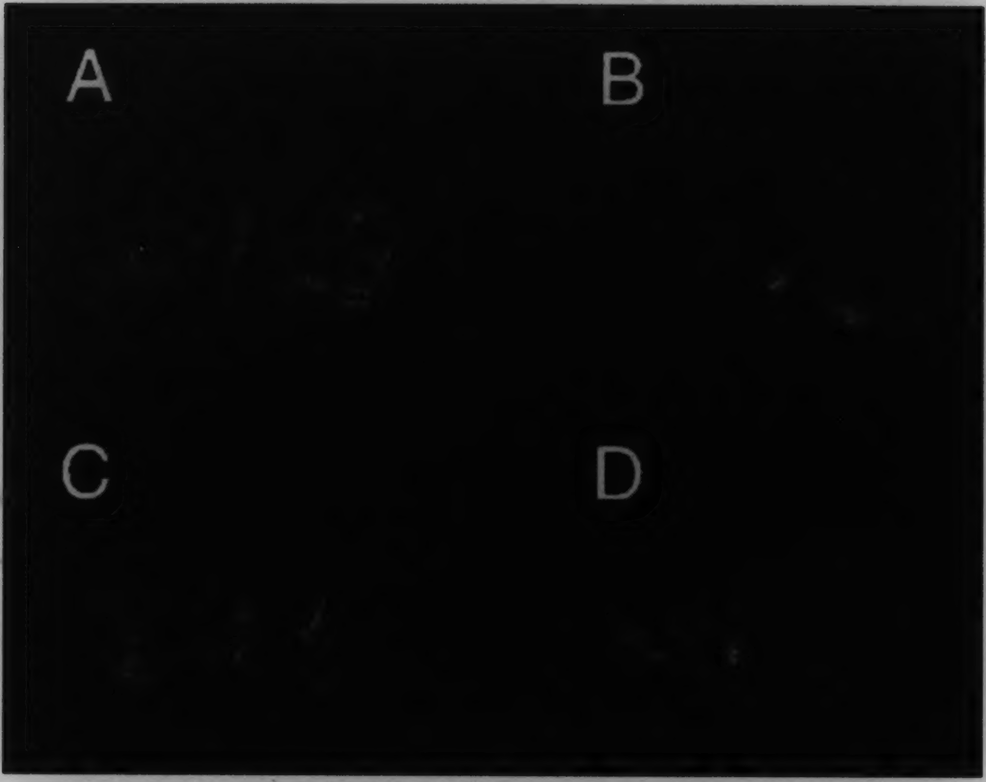


Figure 10- *Molecular surface of A. bound crystal structure of microcystin-LR, B. docked average solution structure of motuporin, C. docked crystal structure of okadaic acid, and D. docked crystal structure of calyculin A calculated by GRASP (Ref. 33). The surface is colored on the basis of distance to the surface of PP-1c. White represents the surface areas of the toxins close to PP-1c (0-1.5Å), blue corresponds to intermediate proximity (1.5Å), and red relates to distant (over 3Å) proximity.*

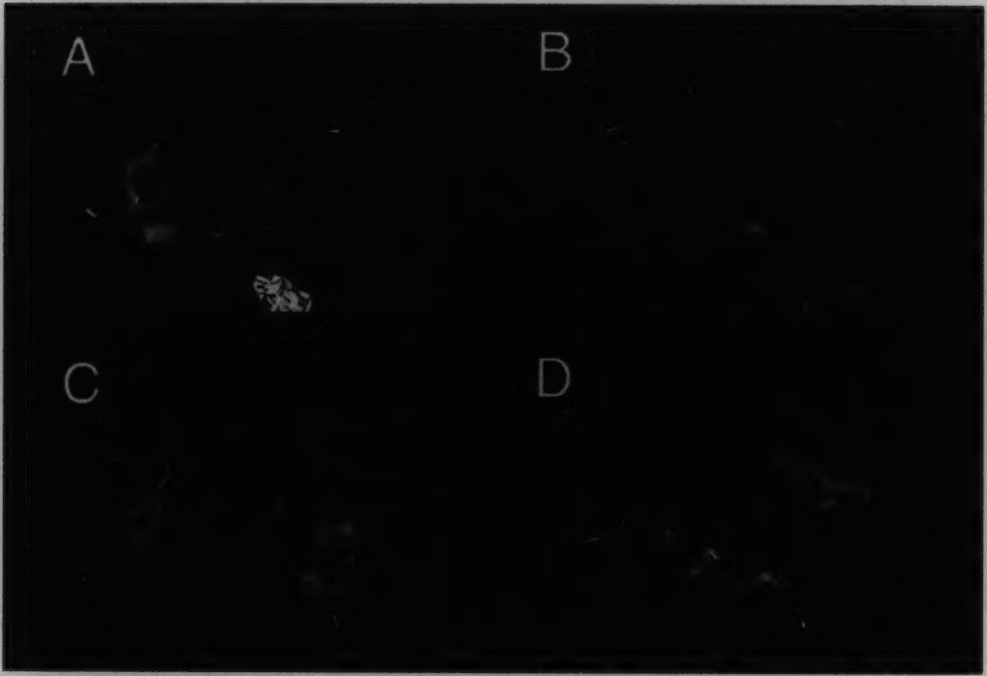


Figure 11- Molecular surface of PP-1c displaying the surface area where the toxins bind as calculated by GRASP (Ref. 33). White represents the surface areas of PP-1c close to the toxin (0-1.5Å), blue corresponds to intermediate proximity (1.5Å), and red relates to distant (over 3Å) proximity. A. is the bound crystal of microcystin-LR, B. is docked average solution structure of motuporin, C. is the docked crystal structure of okadaic acid, and D. is the docked crystal structure of calyculin A.

Discussion

Microcystin-LR is a potent inhibitor of PP-1c, with a high affinity for the enzyme (K_I in the low nanomolar range). With the exception of minor differences noted in the **Results** section, the free NMR solution structure of microcystin-LR matches its crystal structure when complexed with PP-1c (Fig. 2). Can we understand the reasons for why microcystin-LR would not change? Retaining the cyclic backbone of microcystin-LR would reduce the unfavorable conformational entropy lost upon binding[2]. The rigid saddle shaped backbone provides the proper framework for the hydrophobic, ionic, and covalent interactions between microcystin-LR and the phosphatase. These interactions involve the Adda and Masp/Glu residues which are essential for inhibition[9,13,21] and the Mdha residue which is required for the secondary covalent linkage.

The rare, hydrophobic, and long Adda amino acid located behind the rigid saddle is flexible in solution and therefore is able to adjust to fit into a hydrophobic groove of PP-1c[2]. Because hydrophobicity probably is the initial driving force behind the binding of microcystin-LR to PP-1c, the Adda residue is responsible for anchoring the cyclic backbone ring into its bound position. With respect to the Adda residue, the hydrophobic interactions probably more than makes up for the conformational entropy cost that is required to stabilize the Adda side chain into one conformation in the bound state. The other contributing residues for inhibition, the Masp and Glu, have D- oriented negatively charged carboxyl groups located underneath the saddle which are both oriented to interact with the positively charged Arg 96 of PP-1c. Again, the D- nature and saddle shape backbone allow these carboxyls to be in a position to interact with the Arg 96. Finally, the secondary covalent linkage that forms after inhibition (does not contribute to inhibition) is dependent on the modified amino acid Mdha being located at the front, top of the saddle to covalently link to Cys 273.

There are other examples of peptides, either flexible or rigid by nature, that bind to their target proteins with little change in conformation as we have discussed for microcystin-LR. The free solution structure in water of disulphide bond linked desmopressin[22] (Wang *et al.*, submitted) and cyclic cyclosporin A[23,24,25] were found to be similar to the bound form (with desmopressin one of the five families of the free water structure matches the bound form, while only one free form of cyclosporin A exists in water). In the case of hirudin, a linear peptide inhibitor of the interaction of thrombin with fibrinogen and regulatory proteins in blood coagulation, the COOH-terminal region which binds to thrombin was found to have the same global fold free and bound to thrombin[26]. Oxytocin, a disuphide linked hormonal peptide, has been compared free and bound to its carrier protein neurophysin[27]. The two most important residues for binding, Cys 1 and Tyr 2, do not change when complexed to neurophysin[27]. The rest of the ring undergoes a conformational change, but this is thought to occur after binding takes place[27]. Taken together these examples coupled with our microcystin-LR data indicate that determination of free solution structures in water of many biologically active peptides may provide useful insight into their function since these structures may be retained upon binding to their target(s).

The high degree of similarity between microcystin-LR and the more hydrophobic microcystin-LL (Fig. 3A and Fig. 3B) suggests that the microcystin class of inhibitors probably have a basic saddle shaped backbone with the Adda side chain sticking out behind it. This would account for the similar nature of the microcystin members in binding to PP-1c. Other published microcystin structures have been carried out in the presence of dimethyl sulfoxide (DMSO) solvent. The published microcystin-LR structure in DMSO[28] was found to have 3 different conformational families with a highly compact ring structures that do not appear to match the saddle shape ring of the free solution structure in water[1] or bound crystal structure[2]. The structure of microcystin-LY in DMSO[28] was described as being a boat-like ring structure with the Adda residue protruding form the concave side which is more in line with the microcystin-LR and microcystin-LL structures. Recently, a paper was published comparing the structure of microcystin-LR in water and in DMSO[29]. In this study, microcystin-LR was found to have the same structure in water and DMSO[29] and both structures were similar to the previously published free solution structure[1] and bound crystal structure[2]. Thus, the conformation of microcystin-LR seems to be unclear in DMSO[28,29] but is confirmed in water[1,2,29].

Accurate docking of the free solution structure of microcystin-LL, motuporin (Fig. 5B), okadaic acid (Fig. 7), and calyculin A (Fig. 8) relies upon the assumption that like microcystin-LR their free structures would not dramatically change upon binding. Of course, it is also dependent on the assumption that PP-1c itself does not differ in conformation when microcystin-LR is bound to it verses when microcystin-LL, motuporin, okadaic acid, or calyculin A is bound. The latter assumption is supported by the X-ray crystallographic determination of PP-1c without microcystin-LR which was found to be in overall agreement with the previously published PP-1c crystal structure with microcystin-LR[2,30]. This indicates that these inhibitors probably do not have a signficant impact on the overall tertiary structure of PP-1c when binding to the enzyme. Because microcystin-LL and motuporin are similar to microcystin-LR in terms of structure and function, we believe that the best possible starting point for docking of both inhibitors would be in the position relative to PP-1c where they are superimposed onto bound crystal structure of microcystin-LR. This was also true for okadaic acid and calyculin A that while although chemically different from microcystin-LR their free crystal structures were very similar to the tertiary structure of microcystin-LR.

As mentioned in the Methods, we applied a rigorous test for the docking macro used in this study the minimized average free solution structure of microcystin-LR was docked in the same manner as microcystin-LL and motuporin. The final position of the docked microcystin-LR was only slightly shifted from its starting point (data not shown) meaning that Metropolis Monte Carlo docking was successful in placing the free solution structure of microcystin-LR very near the bound structure as determined by X-ray crystallography[2]. This implies that the docking protocol could do the same for the other inhibitors too.

The similarities in the backbone atoms of Masp, arginine/valine, and Adda backbone atoms, positioning of the Adda side chain, and the Glu and Masp carboxyls between microcystin-LR and motuporin probably are responsible for both cyclic peptides being specific, high affinity, and tight binding inhibitors of PP-1c (Fig. 4). While they are functionally and structurally similar there are differences between the microcystin and nodularin class inhibitors. Microcystins covalently bind to PP-1c while nodularins are unable to covalently link to the enzyme. We proposed that the basis for this functional difference was found in the structural displacement of the Mdha residue (involved in the covalent linkage with PP-1c) in microcystin-LR with its counter part residue Mdhb in motuporin[1]. By comparing the bound crystal structure of microcystin-LR[2] with the free solution structure of motuporin superimposed in the same manner[1] (Fig. 4) we further substantiate our previous hypothesis[1] that a large displacement (7.13Å as seen in Fig. 4) occurs between the beta carbon of the Mdha residue located at the front, top of the saddle in bound microcystin-LR and the beta carbon of the Mdhb residue in motuporin. The

Mdhb residue is located far below the Mdha residue, with it being located in the Leu position of microcystin-LR.

Docking of the motuporin solution structure onto the crystal structure of PP-1c by Metropolis Monte Carlo resulted in only a minor displacement from the starting point of motuporin superimposed onto bound microcystin (Fig. 5A verses Fig. 5B). With docked motuporin the distance between the beta carbon of the Mdhb residue to the sulpur atom on the side chain of Cys 273 where the potential covalent linkage would be is over 10Å. Again, this suggests that the Mdhb residue of motuporin is out of position, being too far from Cys 273 to form a covalent link with PP-1c. Also, while there are some common protein phosphatase residues that are in proximity to both microcystins and motuporin, the relative lack of a front part of saddle in motuporin results in no proximity connections to PP-1c residues 274-6. Viewing this docked motuporin complexed with PP-1c and comparing it to the crystal complex[2] between microcystin-LR and PP-1c (Fig. 6A verses Fig. 6B) shows how the Mdhb residue in motuporin is not only surface accessible but highly exposed to the solvent which is not true for the covalently bonded residue Mdha in microcystin-LR. The inability of motuporin and nodularin to form a covalent link to PP-1c and PP-2Ac may ultimately explain why the nodularins may be able to function as carcinogens. Since these toxins form no covalent linkage to protein phosphatases, their electrophilic Mdhb residues may be free to form direct adducts with nucleophilic groups on other informational macromolecules either when bound to protein phosphatase (as viewed in Fig. 6B) or following a dissociation event from these enzymes. It is conceivable that this may ultimately explain why microcystin-LR appears to be devoid of carcinogenic properties since it would be expected to remain covalently attached to PP-1c/PP-2Ac, and therefore be unable to form an adduct with other macromolecules via the Mdha residue.

The model for docked okadaic acid has distinct similarities to the bound tertiary structure and position for microcystin-LR (Fig. 7A and 7B). This would explain how these diverse toxins are able to have the same function as tumor promoters by inhibiting PP-1c, and agrees with competition studies between okadaic acid and microcystin which show how these toxins bind in a mutually exclusive manner[2,8]. The model is also in agreement with the previous biochemical studies that examined the interaction between okadaic acid and PP-1c by altering either the toxin or the enzyme. For example, O5 of okadaic acid has been altered by the additions of bulky groups which do not affect the binding of the toxin. In the model, this oxygen is highly solvent accessible and is not near the phosphatase-toxin interface. Therefore, any additions to this oxygen would not get in the way of okadaic acid binding to PP-1c. The oxygen atoms at the C1 carboxyl end of okadaic acid (O1, O2, and O3), especially O1, are to some extent situated near Tyr 272, Val 250, Arg 96, and Arg 221. Methylation of O1 has been shown to detrimentally affect okadaic acid binding, which in the model would be explained by the disruption of O1's interaction with one of the 4 residues mentioned in the previous sentence, most probably Tyr 272. Mutations[2,30,31] of PP-1c residues 271-277 that affect the sensitivity of okadaic acid for PP-1c can be explain by the association of the C1 carboxyl terminus with Tyr 272 and that other regions of okadaic acid are near (Fig. 7B) and probably interacting with these phosphatase residues (eg. Phe 276). The O11 atom connected to C27 when modified has reduced okadaic acid activity, and although this atom is somewhat solvent exposed, it is within 4Å of Tyr 134 which may be an important interaction for toxin binding. Other interesting points with the okadaic acid model are that the O10 (C24) hydroxyl group is very near (less than 3Å) to Arg 221. The protruding C25-C41 double bond is at the phosphatase-toxin interface but is far from Cys 273 or any other group likely for covalent linkage so okadaic acid probably doesn't covalently link with PP-1c. Finally, the C14-C15 double bond which needs to be in the E conformation for okadaic acid to be active[32] is to some extent close to Glu 275, but it is conceivable that a change in the configuration of this double bond could alter the overall structure of okadaic acid thus influencing its binding.

The model for calyculin A, like okadaic acid, shares many features with bound microcystin-LR (Fig. 8A and Fig. 8B). Unlike okadaic acid though, not much appears to be known about what parts of calyculin A are important for its activity. In calyculin A there are hydroxyl groups and a carboxyl group that as in the case of okadaic acid could be important for activity. As well, there is a nitrogen triple bonded to a carbon at the hydrophobic end of calyculin A that to some extent is solvent exposed and appears not to be in very close proximity to any PP-1c residue. The phosphate group probably contributes to binding though by interacting with Arg 96 and/or Arg 221, although in the model these are farther apart than the Masp or Glu is to the PP-1c residues in bound microcystin-LR. However, the phosphate group is a stronger negative charge and may not be required to be as close as Masp or Glu to interact with the arginines.

Previously Quinn *et al.*[32] had previously recognized common structural components (namely the acidic group and hydrophobic region) that microcystin-LR, okadaic acid, and calyculin A share, however they did not have an accurate structure of microcystin-LR and the structure of PP-1c to work with. With the current availability of the free solution structures of microcystin-LR, microcystin-LL, and motuporin, as well as the crystal structure of microcystin-LR bound to PP-1c, we have extended these preliminary studies further by docking the previously published free crystal structures of okadaic acid and calyculin A to PP-1c in the same manner as microcystin-LL and motuporin. These docked models of okadaic acid and calyculin A (when compared with the microcystins and motuporin) now provide a compelling hypothesis to account for how these diverse toxins are able to inhibit PP-1c.

References

1. Bagu, J.R., Sönnichsen, F.D., Williams, D., Andersen, R.J., Sykes, B.D., and Holmes, C.F.B. (1995) *Nature Stuctural Biology* **2**, 114-116

2. Goldberg, J., Huang, H., Kwon, Y., Greengard, P., Nairn, A.C., and Kuriyan, J. (1995) *Nature* **376**, 745-753

3. Tachibana, K., Scheuer, P.J., Tsukitani, Y., Kikuchi, H., Van Engen, D., Clardy, J., Gopichand, Y., Schmitz, F.J. (1981) *J. Am. Chem. Soc.* **103**, 2469-71

4. Kato, Y., Fusetani, N., Matsunaga, S., Hashimoto, K., Fujita, S., and Furuya, T.J. (1986) *J. Am. Chem. Soc.* **108**, 2780-1

5. Carmichael, W.W. (1994) *Scientific American* **270(1)**, 78-86

6. Honkanen, R.E., Zwiller, J., Moore, R.E., Daily, S.L., Khatra, B.S., Dukelow, M., and Boynton, A.L.(1990) *J. Biol. Chem.* **265**, 19401-19404

7. Yoshizawa, S., Matsushima, R., Watanabe, M.F., Harada, K., Ichihara, A., Carmichael, W.W., and Fujiki, H. (1990) *J. Cancer Res. Clin. Oncol.* **116**, 609-614

8. MacKintosh, C., Beattie, K.A., Klumpp, S., Cohen, P., and Codd, G.A. (1990) *FEBS Lett.* **264**, 187-192

9. Nishiwaki-Matsushima, R., Nishiwaki, S., Ohta, T., Yoshizawa, S., Suganuma, M., Harada, K., Watanabe, M.F., and Fujiki, H. (1991) *Jpn. J. Cancer Res.* **82**, 993-996

10. Holmes, C.F.B. and Boland, M.P. (1994) *Current Opinion in Structural Biology* **3**, 934-943

11. Eriksson, J.E., Toivola, D., Meriluoto, J.A.O., Codd, G.A., Kass, G.E.N., Karaki, J., Han, Y.G., and Hartshorne, D. (1990) *Biochem. Biophys. Res. Commun.* **173**, 1347-1353

12. Ohta, T., Nishiwaki, R., Yatsunami, J., Komori, A., Suganuma, M., and Fujuki, H. (1992) *Carcinogenesis* **13**, 2443-2447

13. Rhinehart, K.L., Harada, K.I., Namikoshi, M., Chen, C., Harvis, C.A., Munro, M.H.G., Blunt, J.W., Mulligan, P.E., Beasley, V.R., Dahlem, A.M., and Carmichael, W.W. (1988) *J. Am. Chem. Soc.* **110**, 8557-8558

14. Craig, M., McCready, T.L., Luu, H.A., Smillie, M.A., Dubord, P., and Holmes, C.F.B. (1993) *Toxicon* **31(12)**, 1541-1549

15. Ohta, T., Sueoka, E., Iida, N., Komori, A., Suganuma, M., Nishiwaki, R., Tatematsu, M., Kim, S., Carmichael, W.W., and Fujiki, H. (1994) *Cancer Research* **54**, 6402-6406

16. DeSilva, S.D., Williams, D.E., Andersen, R. J., Klix, H., Holmes, C.F.B., and Allen, T.M. (1992) *Tetrahedron Lett.* **33**, 1561-1564

17. MacKintosh, R.W., Dalby, K.N., Campbell, D.G., Cohen, P.T.W., Cohen, P., and MacKintosh, C. (1995) *FEBS Letters* **371**, 236-240

18. Cohen, P., Holmes, C.F.B., and Tsukitani, Y. (1990) *Trends Biochem Sci* **15**, 98-102

19. Fujuki, H. and Suganuma, M. (1993) *Advances in Cancer Research* **61**, 143-194

20. Allen, F.H., Kennard, O., and Taylor, R. (1983) *Acc. Chem. Res.* **16**, 146-153

21. Stotts, R.R., Namikoshi, M., Haschek, W.M., Rinehart, K.L., Carmichael, W.W., Dahlem, A.M., and Beasley, V.R. (1993) *Toxicon* **31**, 783-789

22. Wang, J., Hodges, R.S., and Sykes, B.D. (1995) *J. Am. Chem. Soc.* **117**, 8627-34

23. Wenger, R.M., France, J., Bovermann, G., Walliser, L., Widmer, A., and Widmer, H. (1994) *FEBS Letters* **340**, 255-9

24. Tayar, N.E., Mark, A.E., Vallat, P., Brunne, R.M., Testa, B., and Gunsteren W.F.v. (1993) *J. Med. Chem.* **36**, 3757-64

25. Altschuh, D., Vix, O., Rees, B., and Thierry, J.C. (1992) *Science* **256**, 92-94

26. Ni, F., Ripoll, D.R., and Purisima, E.O. (1992) *Biochemistry* **31**, 2545-2554

27. Rose, J.P., Wu, C., Hsiao, C., Breslow, E., and Wang, B. (1996) *Nature Structural Biology* **3**, 163-9

28. Rudolph-Bohner, S., Mierke, D.F., and Moroder, L. (1994) *FEBS Letters* **349**, 319-323

29. Trogen, G.B., Annila, A., Eriksson, J., Kontteli, M., Meriluoto, J., Sethson, I., Zdunek, J., and Edlund, U. (1996) *Biochemistry* **35**, 3197-3205

30. Egloff, M., Cohen, P.T.W., Reinemer, P., and Barford, D. (1995) *J. Mol. Biol.* **254**, 942-959

31. Zhang, Z., Zhao, S., Long, F., Zhang, L., Bai, G., Shima, H., Nagao, M., and Lee, E.Y.C. (1994) *J. Biol. Chem.* **269**, 16997-17000

32. Quinn, R.J., Cherie, T., Suganuma, M., and Fujiki, H. (1993) *Bioorganic & Medicinal Chemistry Letters* **3**, 1029-34

33. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Prot. Struct. Funct. Genet.* **11**, 281-96.

Thesis Chapter 4

Use of structure-based thermodynamics in the analysis of protein phosphatase-1 and microcystin-LR and -LL docked complexes

Pages 69-102 of this chapter submitted to Protein Science as Lavigne, P., Bagu, J.R., Boyko, R., Willard, L., Holmes, C.F.B., and Sykes, B.D. (Submitted).

Pages 103-157 of this chapter non-published work primarily by Bagu, J.R..

Summary

The relationship between the structure of a free ligand in solution and the structure of its bound form in a complex is of great importance to the understanding of the energetics and mechanism of molecular recognition and complex formation (dissociation). In this study, we use a structure-based thermodynamic approach to study the dissociation of the complex between the toxin microcystin-LR (MLR) and the catalytic domain of protein phosphatase-1 (PP-1c) for the crystal structure is known. We have calculated the thermodynamic parameters (enthalpy, entropy, heat capacity and free energy) for the dissociation of the complex from its X-ray structure and found the calculated dissociation constant ($4.05 \cdot 10^{-11}$) to be in excellent agreement with the reported inhibitory constant ($3.86 \cdot 10^{-11}$). We have also calculated the thermodynamic parameters for the dissociation of 47 PP-1c:MLR complexes generated by docking an ensemble of NMR solution structures of MLR onto the crystal structure of PP-1c. In general, we observe that the lower the r.m.s.d. of the docked complex (compared to the X-ray complex) the closer its free energy of dissociation (ΔG°_d) is to that calculated from the X-ray complex. On the other hand, we note a significant scatter between the ΔG°_d and the r.m.s.d. of the docked complexes. We have identified a group of seven docked complexes with ΔG°_d very close the one calculated from the X-ray complex but with significantly dissimilar structures. The structural analysis of these docked complexes reveals that the specific interactions are made correctly but that significant differences in the conformation of MLR still exist. The analysis of the corresponding enthalpy and entropy of dissociation shows a compensation effect suggesting that MLR molecules with significant structural variabilty can bind PP-1c and that substantial conformational flexibility in the PP-1c:MLR complex may exist in solution.

Introduction

Protein phosphorylation is a general mechanism for the regulation of many important cellular processes[1-3]. Due to the reversible nature of phosphorylation there is generally an antagonistic relationship between activation of cellular processes achieved by protein kinases, and deactivation of cell signals by protein phosphatases. Dephosphorylation of serine and threonine is mainly accomplished by 4 subgroups of phosphatases; protein phosphatase-1 (PP-1c), -2a, -2b (calcineurin), and -2c[1]. Some of these classes (PP-1c and PP-2a) are inhibited by metabolites of cyanobacteria (*e.g.* microcystins and nodularins), dinoflagellates (*e.g.* okadaic acid), and compounds isolated in sponges (*e.g.* calyculin A). These metabolites are liver toxins and have powerful tumor promotion activity linked to morphological changes in hepatocytes[4]. The microcystins are cyclic 7 amino acid peptides containing several unusual amino acids. There are several varieties of microcystins with the differences normally localized to changes in the two variable amino acids or alterations in the methylaspartic acid and/or N-methyldehydroalanine residues[5]. Two examples are microcystin-LR (MLR) and microcystin-LL (MLL), wherein Arg has been substituted for Leu. Microcystins are able to covalently link to PP-1c with the N-methyldehydroalanine residue joining with Cys 273[6,7]. This covalent linkage is time dependant and has no impact on the initial inhibition of PP-1c or PP-2a. The consequence of the covalent linkage is to irreversibly inhibit the phosphatase preventing any further activity.

The NMR solution structure of MLR has been determined[8] as well as the structure of PP-1c covalently complexed with MLR[9]. The free and bound forms of MLR were found to have similar overall structures and, most strikingly, the conformation of the cyclic backbone of the solution structure of MLR is almost identical to the structure in the complex. The relationship between the structure of a ligand in its bound form and in its free form is of particular interest to the field of drug design. For example, if one could be able to discover the active conformation of a peptide in a family of NMR structures this would accelerate drug discovery. One way of addressing this problem is to develop approaches that could be capable of calculating reliable energetic or thermodynamic parameters for the association (dissociation) of complexes generated by docking solutions structures onto a target of known structure. Therefore the solution structures of the ligands present in complexes that give calculated thermodynamic parameters that agree well with measured ones should be close to the active form of the ligand or at least competent binding conformations.

In a first attempt to understand the relationship between the free form and the bound form of MLR in the inhibition process of PP-1c, we first generated a model for the PP-1c:MLR complex by a rigid body docking procedure using the average solution structure of MLR[10]. The success of docking the average NMR solution structure to PP-1c in the same position as the bound crystal MLR allowed for further successful dockings of microcystin-LL, motuporin, okadaic acid, and calyculin A. These toxins all had similar three-dimensional structures despite significant primary structural differences and were proposed bind to the same site as MLR. The quality of the models were assessed on the basis of surface complementarity and potential energy obtained from a molecular mechanics force field[10].

While molecular mechanics force fields are useful to maintain proper non-covalent and covalent stereochemistry they seem not to be able to discriminate the correct fold within clusters of docked complexes with minimal potential energies. This problem, sometimes referred to as the "docking problem", involves the discrimination between the correct answer and the "false positives" that have similar potential energy but incorrect structures[11]. Recently, recourse to solvation free energy corrections has been shown to partially solve that problem[11]. On the other hand, empirical free energy functions have been successfully used to reliably calculate the binding free energies (ΔG_{bind}) or relative ΔG_{bind} from the structure of complexes[12]. Therefore, it has been proposed[13,14] that the minimization of accurate or realistic empirical free energy function might alleviate the docking problem. The development and the use of accurate free energy functions is of the outmost importance both in the protein folding problem and molecular recognition fileds. Ultimately, what is needed is an empirical free energy function(s) that is realistic (faithful) or accurate enough to reproduce both the experimental configuration and the corresponding thermodynamics for the folding and binding processes of polypeptide chains.

Binding free energies (ΔG_{bind}) can be computed from first principles using statistical mechanical approaches[15]. Although these calculations rely on exact results of statistical mechanics they are not easily tractable. Empirical free energy functions can obtained from structural database statistics or from linear regressions fitting different empirical free energy parameters (scaling with molecular surface) with experimental

ΔG_{bind}. Both approaches have been used with some success, although the latter approach seems to be restricted to homologous systems[12]. A third approach that is transferable to other systems and that has been successful to predict the experimental ΔG_{bind} (or $\Delta\Delta G_{bind}$ for mutants) separates the total ΔG_{bind} into hydrophobic force (per Å^2 of surface), electrostatic interactions (Poisson-Boltzmann equation), conformational entropy and overall rotational and translational entropy contributions[16,17]. A fourth approach that can calculate the free energy of binding from a parameterization (per Å^2 of polar and non-polar ASA) of the heat capacity, enthalpy and solvation entropy obtained from a global fit of structural and thermodynamic database of globular proteins[18-20]. This method has the advantage of being tractable and of separating the ΔG into enthalpic and entropic contributions. This approach has been successfully used to calculate from the crystal structure of protein-peptide, protein-ligand and protein-protein complexes the energetics of dissociation (enthalpy (ΔH), entropy (ΔS), and heat capacity change ΔC_p) of complexes that agree well with experimentally determined values[19,20]. This approach has also been utilized in order to validate the model structure of a complex[21].

In this study we use a slightly modified version of the structure based approach described above to address the relationship between the bound conformation and the free solution structure of MLR. More specifically, we calculate the thermodynamic parameters for the dissociation of 47 complexes generated by docking an ensemble of NMR solution structures of MLR on the crystal structure of the PP-1c. We compare these values with the ones calculated form the crystal structure of the complex and explore the relationship between structural diversity and the energetics of the different complexes.

Methods

Docking Procedure

Metropolis Monte Carlo docking of the 46 calculated solution structures and the average minimized solution structure of MLR to the crystal structure of PP-1c was accomplished using the Monte Carlo macro in Insight II version 2.3. This technique proved useful in docking other toxins as well[10]. Based on previous work[10], the starting positions of all calculated MLR solution structures were determined by superimposing their backbone atoms onto the backbone of bound MLR[9] which was then removed. Docking calculations were performed using 2000 trials at a temperature of 50 K[10]. The forcefield used was CVFF[22].

This docking technique involves the minimization of the Van der Waals and Coulomb potential energies between two rigid bodies by altering their relative positions (in this case the MLR and PP-1c structures). The new state is rejected or accepted based on the new potential energies. Normally accepted states have reduced energies, however, the docking procedure allows higher energy states to be accepted on occasion in order that the docked structures are not trapped in a local energy minimum. However, the jump to higher energy states that are accepted become reduced as the number of docking iterations increases. By about 2000 trials accepted states at higher energy than the previous accepted state have only minor increases in energy.

Structure-based free energy calculations

The free energy of dissociation of a complex, $\Delta G^{\circ}_{d}(T)$, is classically given by the following equations:

$$\Delta G^{\circ}_{d}(T) = -RT \cdot \ln K_{d}, \qquad (1)$$

$$\Delta G^{\circ}_{d}(T) = \Delta H^{\circ}_{d}(T) - T \cdot \Delta S^{\circ}_{d}(T), \qquad (2)$$

where $\Delta H^{\circ}_{d}(T)$ is the temperature dependent standard enthalpy of dissociation and $\Delta S^{\circ}_{d}(T)$ is the temperature dependent standard entropy of dissociation.

The calculation of the $\Delta H^{\circ}_{d}(T)$

The following description addresses the structural parameterization of protein unfolding enthalpy calculation. Since protein unfolding and protein dissociation are governed by the same molecular forces, the parameterization developed for protein unfolding is assumed to apply to protein dissociation[23] throughout the text.

$\Delta H^{\circ}_{d}(T)$ is given by:

$$\Delta H^{\circ}_{d}(T) = \Delta H^{\circ}_{d}(T^{\circ}) + \Delta C_{p} \cdot (T - T^{\circ}), \tag{3}$$

where $\Delta H^{\circ}_{d}(T^{\circ})$ is a standard reference enthalpy of dissociation at some reference temperature (T°). $\Delta C_{p}(T)$ is the temperature independent heat capacity change upon dissociation of the complex. It is a good approximation to consider ΔC_{p} temperature independent from 0 to 85°C[23].

The experimental ΔC_{p} of unfolding of a series proteins for which the both the thermodynamics of unfolding and the crystal structure were well characterized can be reliably calculated by a linear combination of the change in solvent ASA of polar (ΔASA_{pol}) and non-polar (ΔASA_{np}) atoms through the following empirical relationship[18]:

$$\Delta C_{p} = 0.45 \cdot \Delta ASA_{np} - 0.26 \cdot \Delta ASA_{pol}, \tag{4}$$

where the parameters have units of $cal \cdot K^{-1} \cdot mol^{-1}$ $Å^{-2}$. ΔC_{p} of protein unfolding has been shown to come mainly from the hydration of atoms that become exposed to the solvent upon unfolding and as can be seen in the preceding relationship the hydration polar atoms and non-polar atoms have opposite contributions[18,23].

Similarly it has been shown that the experimental unfolding enthalpy $\Delta H^{\circ}_{d}(T)$ at 60°C of the same series of proteins could be calculated, within 6% error, with the following empirical rule[24,25]:

$$\Delta H^{\circ}_{d}(60) = 31.4 \cdot \Delta ASA_{pol} - 8.44 \cdot \Delta ASA_{np}, \tag{5}$$

where the parameters have units of cal · mol^{-1} · Å^{-2}. It is implicit that this empirical function accounts for the change in enthalpy resulting for breaking non-covalent bond (H-bonds, salt bridges and van der Waals interactions etc.) and solvating these atoms[25]. Other contributions such as protons transfer have to be accounted for if they are coupled to the unfolding or dissociation process[21,23].

Therefore once the changes in ASA for a dissociation are calculated, a corresponding $\Delta H^{\circ}_{d}(T)$ can be calculated by combining Equations 4 and 5:

$$\Delta H^{\circ}_{d}(T) = \Delta H^{\circ}_{d}(60) + \Delta C_{p} \cdot (T - 60), \tag{6}$$

The calculation of the $\Delta S^{\circ}_{d}(T)$

In absence of proton transfer, the standard entropy of dissociation, ΔS°_{d} can be assumed to correspond to the sum of the changes in solvation entropy ($\Delta S_{sol}(T)$), conformational entropy (ΔS_{conf}) and overall rotational/translational entropy (ΔS_{rt}) to account for the appearance of an additional kinetic unit upon dissociation:

$$\Delta S^{\circ}_{d}(T) = \Delta S_{sol}(T) + \Delta S_{conf} + \Delta S_{rt} \tag{7}$$

Of the three contribution, only ΔS_{sol} is assumed temperature dependent and can be broken into contributions arising from the change in solvation entropy resulting from the solvation of polar and non-polar atoms that become exposed upon dissociation. It has been shown that the solvation entropy of non-polar atoms is zero at 112°C[26] and that the temperature at which solvation entropy of polar atoms is equals zero is close to 62°C[27].

Therefore, it has been proposed that $\Delta S_{sol}(T)$[27] can be parameterized by the following relationship:

$$\Delta S_{sol}(T) = 0.45 \cdot \Delta ASA_{np} \cdot \ln(T/384.15) - 0.26 \cdot \Delta ASA_{pol} \cdot \ln(T/335.15) \quad (8)$$

Where the coefficient 0.45 and -0.26 are the ones described in Equation 4.

Murphy et al.[28] proposed the following scheme to account for the change in conformational entropy (ΔS_{conf}) for protein dissociation:

$$\Delta S_{conf} = \Delta S_{bu\rightarrow ex} + \Delta S_{ex\rightarrow u} + \Delta S_{bb} \quad (9)$$

where $\Delta S_{bu\rightarrow ex}$ is the gain in conformational entropy of a side-chain when it becomes exposed after disruption of tertiary or quaternary interactions, $\Delta S_{ex\rightarrow u}$ is the change in conformational entropy of the side-chain when the secondary structure unfolds and ΔS_{bb} is the gain in conformational entropy from the backbone upon unfolding. $\Delta S_{bu\rightarrow ex}$, $\Delta S_{ex\rightarrow u}$ and ΔS_{bb} values for all the amino acids have been estimated from a statistical mechanical analysis[27,29]. In our calculations we use the values reported by D'Aquino et al.[27].

In the present case, it is assumed that no conformational change in the backbones of the PP-1c and MLR occur upon dissociation. This is supported by the fact that MLR is a cyclic peptide and our earlier findings that the conformation of the cyclic backbone of the solution structure of MLR and MLR are identical to that of the bound state of MLR in the crystal structure of the PP-1c:MLR complex[8-10]. The crystal structure of the free PP-1c was shown to be almost identical to that of the complexed form[9,30] indicating that there should not be any major conformational change occurring in the free form of the enzyme. Therefore, the change in conformational entropy of dissociation of PP-1c is assumed to originate solely from the gain in conformational entropy of the side-chains that become exposed upon dissociation ($\Delta S_{bu\rightarrow ex}$) and is scaled as the fraction of the total ASA of the side-chain that is gained and computed according to the following equation:

$$\Delta S_{conf} = \sum_i \frac{\Delta ASA_i}{ASA_i} \cdot \Delta S_{bu\rightarrow ex} \quad (10)$$

where ΔASA_i is the change in ASA of the side-chain of residue i and ASA_i is the ASA of the corresponding side-chain in a fully exposed state. Here we used the ASA values reported by Miller et al.[31]. Special care had to be taken for the conformational entropy and the ASA of the non-natural side-chains of the toxins (see Fig. 1). For the Adda side-chain, we used the empirical equation proposed by Bardi et al.[32] for non natural peptidyl side-chains which relates the number of rotable bonds (1.76 cal · K^{-1} · mol^{-1} per rotable bond) and the number of atoms (0.414 cal · K^{-1} · mol^{-1} per atom) to correct for excluded volume effects[32]. A value of 7.8 cal · K^{-1} · mol^{-1} is calculated for the Adda side-chain. The ASA of Adda side-chain was calculated in the free toxins and amounts to 387 $Å^2$. The other side-chains that had to be ascribed with a ΔS_{conf} where the D-Glu and Masp residues (see Fig. 1). As can be seen, these side-chains consists in a single carboxylate which can rotate around their Cα(sp3)-COO(sp2). As discussed by Pickett & Sternberg[33], the COO^- has a symmetry number of 2 leading to distinguishable rotamers on only 180°. We approximated that the conformational entropy in the buried state is R·ln2 rather than 0 (R·ln1). We assume the change in the conformational entropy of the D-Glu and Masp side-chains to be equal to 0.8 cal · K^{-1} · mol^{-1} (-R·[ln2 - ln3]) by supposing that three distinguishable rotamers can be adopted by the side-chain when free to rotate in the free form of the toxins. We also measured a ASA of 90 $Å^2$ for the side-chains.

The gain in translational and rotational entropy (ΔS_{rt}) seems to be well accounted for by to the cratic entropy[23,28]. The cratic entropy is equal to R ln (1/55) where the ratio is the mole fraction of the additional particle appearing (mixing ideally) upon dissociation at a fictitious 1M standard state in water[34]. The cratic entropy amounts to 8 cal · K^{-1} · mol^{-1} or 2.4 kcal · mol^{-1} at 25°C. On the other hand, the use of this value and its physical basis is a matter controversy[15,35]. However, recent experimental evidence[36,37] and theoretical arguments[38] indicate that the loss in rotational and translational entropy is numerically close to the cratic entropy. Therefore we are also using a value of 8 cal · K^{-1} · mol^{-1} to account for the ΔS_{rt}.

The STC program suite

In order to perform the free energy calculations from the structure of the different complexes, we developed a suite of programs called STC (Structure-based

Thermodynamics calculation). In essence STC consists of two modules. The first module, *CALCASA*, calculates the change in ASA for the dissociation process from the coordinate files in the Brookhaven protein data bank (pdb) format using the algorithm ANAREA[39] as implemented in the program VADAR[40]. The output files consists in the tabulated ASA of every atom of the complex and of both the free forms of the enzyme and the ligand as well the difference in ASA for each atom. The total changes in non-polar (all carbon atoms and sulfur atoms) and polar (all oxygen and nitrogen atoms) are summed up. In addition, the atomic ΔASA are regrouped per residues (and per side-chain) for the calculation of ΔS_{conf} as described above in the next module.

The module *THERMO* calculates the energetics from the ΔASA. From the total changes in ΔASA_{np} and ΔASA_{pol}, the contribution of non-polar and polar atoms to ΔC_p and $\Delta H^o_d(60)$ are calculated. Then according to Equation 6, the ΔH^o_d at the desired T is calculated. In the present study all the calculations are done at 25°C. Similarly, using the proper reference temperatures, the ΔS_{sol} is extrapolated at 25°C using Equation 8. From the ΔASA of the different side-chains involved in the dissociation, the conformational entropy gained for the ligand and the enzyme is calculated with Equation 7. The total entropy change is then taken to be the sum of all the entropic contributions listed in Equation 7. A $\Delta G^o_d(25)$ is calculated using Equation 1.

Results and Discussion

Calculation of the thermodynamics of dissociation of MLR and PP-1c from the X-ray complex

We present on Figure 1 (*top*) the X-ray structure of MLR bound to PP-1c[9], referred to as X-ray MLR in the rest of the text, and (*bottom*) the ensemble of NMR solution structures of MLR determined by NMR[8] and referred to as NMR MLR from now on. Most of the residues in the 7 amino acid cyclic peptide are unique or altered amino acids. Starting at D-alanine (Ala) the sequence is L-leucine (Leu), β-linked D-erythro-β-methylaspartic acid (Masp), L-arginine (Arg), β-[2S, 3S, 8S, 9S]-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid (Adda), γ-linked D-glutamic acid (D-Glu), and N-methyldehydro-alanine (Mdha). The cyclic backbone is saddle shaped[8] with the Arg pointing above the saddle, the large hydrophobic side-chain Adda pointing behind the saddle, and the negatively charged carboyxl groups located underneath the saddle (as observed in the orientation of Fig. 1). The Mdha residue which covalently links with PP-1c is located at the top, front of the saddle. The Leu side-chain is brown, Masp is yellow, Arg is red, Adda is purple, D-Glu is green, and the cyclic backbone is in blue. The backbone of the average solution structure is almost identical to that of the bound form with r.m.s.d. of 0.65Å. It is notable that in the free solution structure the Arg and Adda side-chains are highly flexible adopting multiple conformations. Conversely, the side-chains of the D-Glu and the Masp residues are constrained by the backbone as they have only one rotable bound.

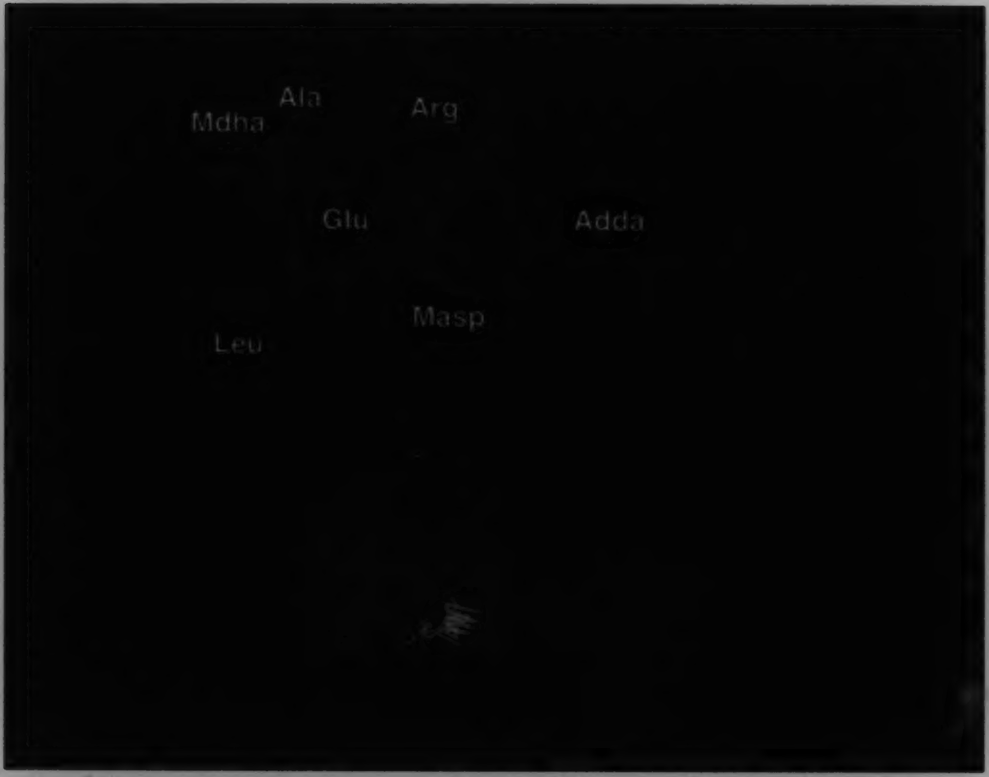


Figure 1- Top: Bound X-ray structure of microcystin-LR (Goldberg et al, 1995). The residues are D-alanine (D-Ala), L-Leucine (Leu), β-linked D-erythro-β-methylaspartic acid (Masp), L-arginine (Arg), β-[2S, 3S, 8S, 9S]-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid (Adda), γ-linked D-glutamic acid (D-Glu), and N-methyldehydro-alanine (Mdha). The cyclic backbone is saddle shaped[8]. ***Bottom:*** *Ensemble of 46 calculated solution structures and average minimized solution structure of MLR[8]. The Leu side-chain is brown, Masp is yellow, Arg is red, Adda is purple, D-Glu is green, and the cyclic backbone is in blue. When compared with the average free solution structure the backbone is almost identical to the bound X-ray form (backbone r.m.s.d. is 0.65Å). For clarity hydrogens are not shown.*

The X-ray complex of MLR bound to PP-1c is represented in Figure 2[9] where only PP-1c residues within 4 Å of MLR are labeled. As discussed elsewhere[9], specific salt-bridges or H-bonds between the MLR and PP-1c involve Masp (Arg 96 and Tyr 134) and D-Glu (Arg 96 and Arg 221). The Arg side-chain of the toxin is also found to lie between the carboxylates of Asp 220 and Glu 275 at the surface the enzyme with distances < 6.0Å giving rise to potential solvent exposed salt-bridges and/or H-bonds. The hydrophobic Adda side-chain interacts with hydrophobic residues lining the hydrophobic groove on the enzyme as described before[9]. Finally, the toxin is found to be covalently linked to Cys 273 (Sγ) through the Cβ of the Mdha. This covalent attachment has been observed to be slow (hours) and not to impair or affect the initial inhibitory action of the toxin.

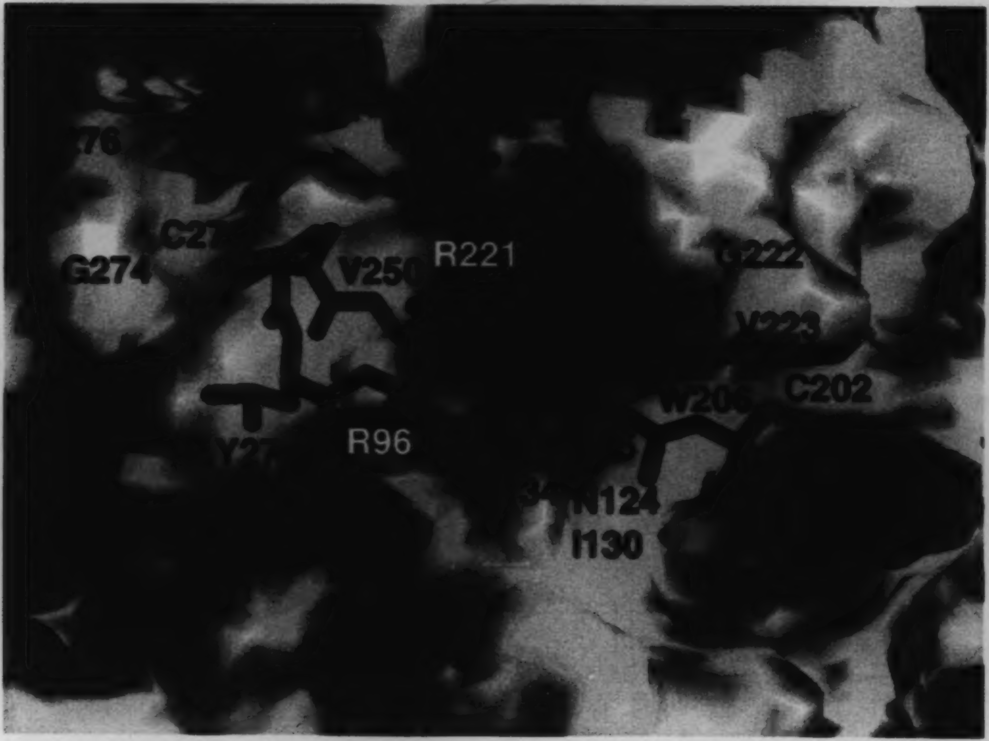


Figure 2- The X-ray complex of MLR bound to PP-1c[9] displayed using the program Grasp[41]. Only PP-1c residues within 4 Å of MLR are labeled. Red are surfaces that are potentially negatively charged, blue are potentially positively charged, yellow represents hydrophobic surfaces, and white are polar surfaces.

The changes in ASA per residue (non-polar and polar) upon dissociation of the X-ray complex are displayed in Figure 3A and the corresponding ΔG°_d on Figure 3B. Overall, MLR exposes 542 Å^2 and 150 Å^2 of non-polar and polar ASA respectively while PP-1c exposes 282 Å^2 of non-polar and 296 Å^2 polar ASA upon dissociation.

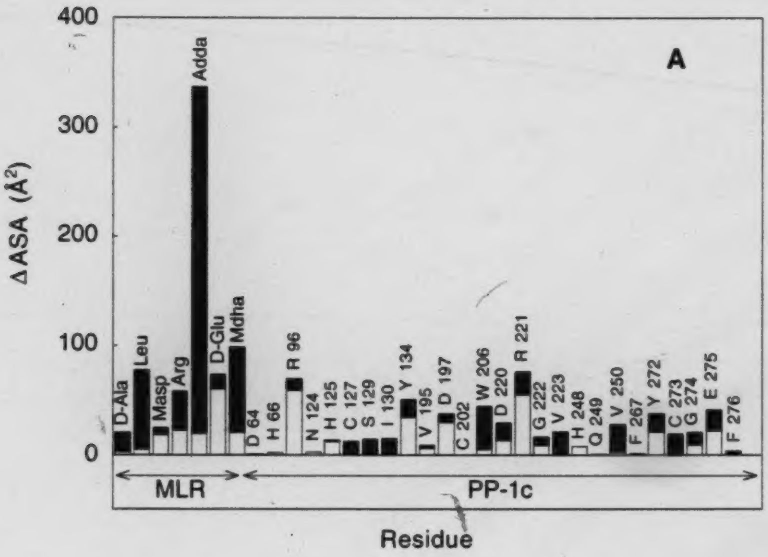


Figure 3A- Changes in ASA per residue (non-polar: black and polar: white) upon dissociation of the complex crystal structure.

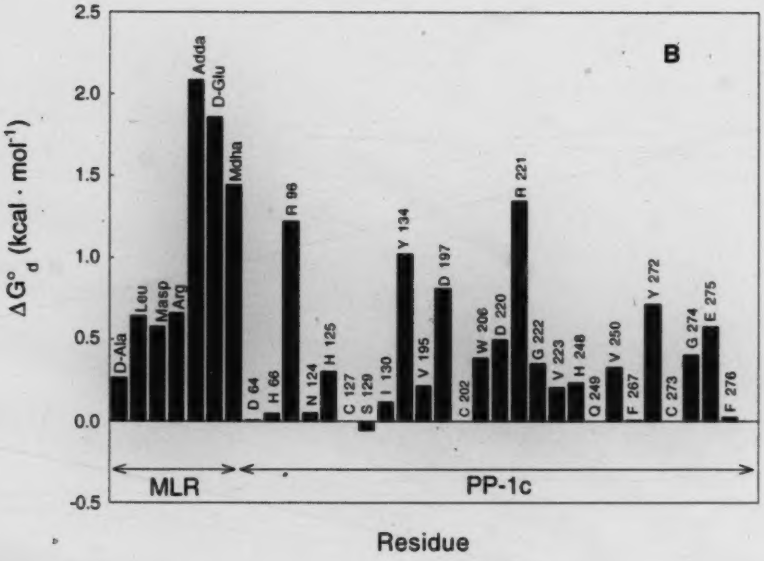


Figure 3B- *Corresponding ΔG°_{d}. MLR exposes 542 $Å^2$ and 150 $Å^2$ of non-polar and polar ASA respectively while PP-1c exposes 282 $Å^2$ of non-polar and 296 $Å^2$ polar ASA upon dissociation.*

Listed in Table 1 are the thermodynamic parameters obtained for the total and per MLR residue ΔASA values for the dissociation of the X-ray complex. As one can notice (Table 1), the unfavorable dissociation free energy calculated: $\Delta G^{o}_{d}(25) = 14.1$ kcal · mol^{-1} is dominated by an overall unfavorable dissociation entropy ($-T\cdot\Delta S^{o}_{d}(25) = 16.0$ kcal · mol^{-1}) with a slightly favorable dissociation enthalpy ($\Delta H^{o}_{d}(25) = -1.87$ kcal · mol^{-1}). It is evident that most of the affinity comes from the unfavorable change in solvation entropy ($-T\cdot\Delta S_{sol}(25) = 21.9$ kcal · mol^{-1}) upon dissociation of the X-ray complex. The Adda residue that buries an extensive amount of non-polar surface, roughly 40 % (317 $Å^2$) of the total ΔASA_{np}, is a major contributor to the affinity (Figure 3B) through the hydrophobic effect, *i.e.* decrease in the solvent entropy upon dissociation ($-T\cdot\Delta S_{sol}(25) = 10.7$ kcal · mol^{-1}, see Table 1). It can be seen that the $T\cdot\Delta S_{sol}(25)$ of the Adda is offset by a favorable $\Delta H^{\circ}_{d}(25)$ and $-T\cdot\Delta S_{conf}$ (Table 1). This side-chain is hypothesized to be critical for the binding of MLR to PP-1c. Indeed, a structural isomer of the Adda side-chain was determined to inhibit PP-1c activity 100 times more weakly than the maternal

MLR[42]. The second largest ΔG°_d (25) on MLR is D-Glu which is dominated by an unfavorable ΔH°_d(25) indicating of favorable interactions with Arg-96 and Arg-221 (Table 1). Interestingly it has been shown that esterification of the D-Glu side-chain has a significant reduction in toxcity suggesting that it is indeed important for the binding of MLR to PP-1c[43]. It is noted on Figure 3B that residues Arg-96, Arg-221 also have a high individual ΔG°_d(25) (> 1 kcal · mol^{-1}) indicating that they are contributing substantially to the affinity of the complex. Interestingly, mutagenesis studies have shown that the replacement of Arg-221 (Arg221Ser) and Arg-96 (Arg96Ala) resulted in drastic reduction in K_i by MLR[44] supporting their high ΔG°_d (25) calculated here.

Table 1. Thermodynamic parameters calculated for the dissociation of PP-1C:MLR X-ray complex and individual residues of MLR[a]

	ΔASA_{ap} (Å^2)	ΔASA_{pol} (Å^2)	ΔC_p kcal · mol^{-1}·K^{-1}	ΔH°_d kcal · mol^{-1}	$-T\cdot\Delta S^\circ_d$[b] kcal · mol^{-1}	$-T\cdot\Delta S^\circ_{sol}$ kcal · mol^{-1}	$-T\cdot\Delta S_{conf}$ kcal · mol^{-1}	ΔG°_d kcal · mol^{-1}	K_{Da}
X-Ray Complex	824	446	0.255	-1.87	16.0	21.9	-5.89	14.1	4.05 *e*-11
D-Ala	18	3	0.0074	-0.32	0.60	0.60	0.000	0.27	-
Leu	72	5	0.0313	-1.54	2.20	2.45	-0.257	0.65	-
Masp	6	18	0.0000	0.60	-0.02	0.05	-0.065	0.58	-
Arg	36	23	0.0102	0.05	0.61	1.02	-0.413	0.66	-
Adda	317	20	0.1376	-6.87	8.95	10.71	-1.761	2.09	-
D-Glu	14	60	0.0000	2.11	-0.26	-0.08	-0.179	1.86	-
Mdha	78	21	0.0297	-1.05	2.50	2.50	0.000	1.45	-

[a] Temperature = 25°C

[b] The $-T\cdot\Delta S^\circ_d$ for the X-ray complex contains contribution ΔS_{tr} but not the $-T\cdot\Delta S^\circ_d$ of individual residues of MLR

As discussed in Methods and stated above, the calculated gain in conformational entropy results solely from side-chains becoming exposed. It is worth pointing out, though, that if the MLR was unfolding (linear) upon dissociation, then the gain in entropy of the backbone would make the affinity of MLR to PP-1c much lower. Assuming a median value for the conformational entropy change for the backbone of ΔS_{bb} ~ 6 cal · K^{-1} mol^{-1} per residue[45], an additional gain in ΔS_{conf} of ~ 42 cal · K^{-1} mol^{-1} can be estimated (-T·ΔS_{conf}(25) ~ -1.5 kcal · mol^{-1} per residue) which would lower the ΔG^o_d(25) from 14.1 kcal · mol^{-1} to practically 0 kcal · mol^{-1}. This is a rough and probably underestimated value since most of the residue of the toxin have more than two rotable bonds and reinforces the idea that the cyclic and folded nature of the backbone of the toxin contributes a lot to its high affinity.

The overall $\Delta H^{o}_{d}(25)$ of –1.87 kcal · mol^{-1} indicates that the disruption of the favorable non-covalent interactions at the interface of the complex is compensated by an almost equally favorable solvation enthalpy of the overall molecular surfaces exposed upon dissociation. It has to be noticed that other factors that could contribute to the enthalpy of dissociation (and the entropy) like putative proton transfer(s)[23] are not treated explicitly here. No change in pK_a of ionizable groups in the catalytic domain or for MLR has been reported. Moreover, no experimental enthalpy (or entropy) of dissociation is available so far to allow us to compare the enthalpy (and entropy) of dissociation calculated with the parameterization used here. In addition, no experimental and conventional K_d is available because of the very high affinity of PP-1c for MLR[46]. On the other hand, it is interesting to notice that the K_d of $4.05 \cdot 10^{-11}$ calculated (Table 1) is excellent agreement with the K_i of $3.86 \cdot 10^{-11}$ reported by Takai et al.[46]. This agreement suggests that the present parameterization satisfactorily describe the energetics of the dissociation of MLR and PP-1c and also suggest that all the assumptions made above appear to be justified.

Assessment of the docked complexes obtained from an ensemble of NMR solution structure of MLR

As mentioned in the Introduction, the development of methods to obtain information about the "bound structure" or binding competent configurations from the structure of the free ligand in solution is very important for our understanding of association (dissociation) reactions and for the rational design of ligands of pharmaceutical interest. We explore in this section the potential use of the structure-based approach to address these issues by analyzing the dissociation thermodynamic parameters obtained for an ensemble of complexes generated from the docking of 47 solution NMR structures of MLR (NMR MLR) onto the crystal structure of PP-1c (X-ray PP-1c) as described in Methods.

Figure 4 (top equilibrium) sketches the process for the structure-based thermodynamic calculations from the X-ray complex. First, the ASA of every atom of the X-ray complex, the dissociated PP-1c and MLR are calculated. Second, the differences in ASA (ΔASA) for every atom of the PP-1c and MLR in both forms are obtained. Finally, the corresponding ΔG°_{d} are computed on a per residue basis as described in Methods.

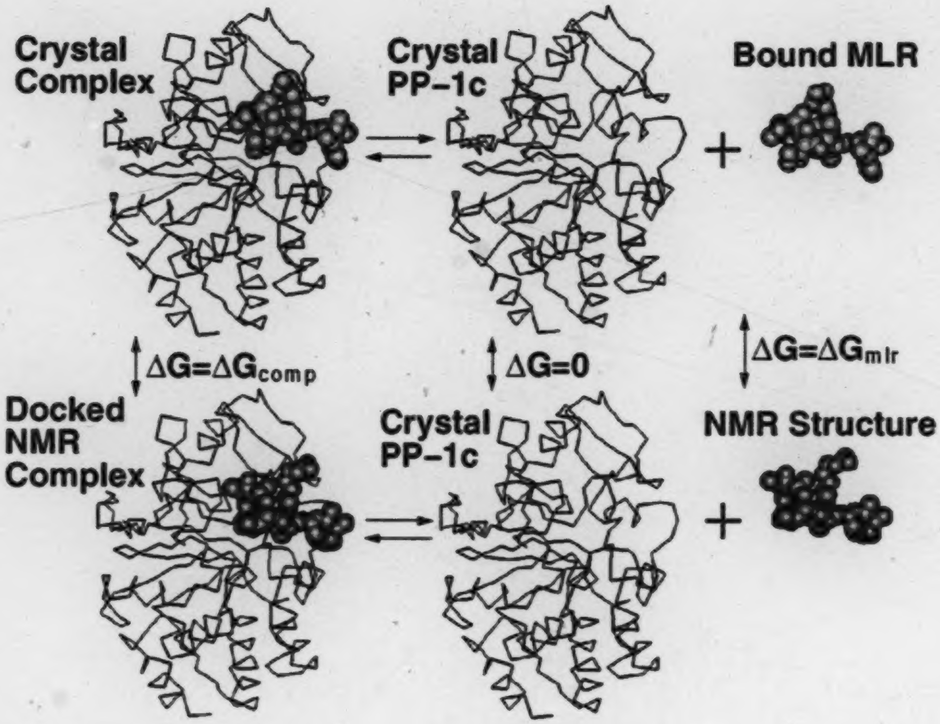


Figure 4- *Schematic representation of the dissociation equilibrium for the X-ray complex (**top**) and the different docked complexes (**bottom**). Since the different docked complexes are different from the X-ray complex, a relative difference in free energy (ΔG°_{comp}), enthalpy (ΔH°_{comp}) and entropy (ΔS°_{comp}) can be calculated for every docked complex from the differences in ASA: $ASA_{docked\ complexes} - ASA_{X\text{-}ray\ complex}$. Similarly one can obtain ΔG°_{MLR}, ΔH°_{MLR} and ΔS°_{MLR} between the different NMR MLR and X-ray MLR from differences in ASA ($ASA_{NMR\ MLR} - ASA_{X\text{-}ray\ MLR}$). $\Delta G^\circ_{d.corr}$ ($\Delta H^\circ_{d.corr}$ or $\Delta S^\circ_{d.corr}$) is obtained by subtracting ΔG_{MLR} (ΔH°_{MLR} or ΔS°_{MLR}) from the ΔG°_d (ΔH°_d or ΔS°_d) of the different docked complexes. $\Delta G^\circ_{PP\text{-}1c} = 0$.*

Before presenting our results, we want to stress a few points about the calculation of the energetic parameters of the ensemble of docked complexes. As is evidenced in the bottom equilibrium of Figure 4, the structures of the different docked complexes as well as the structure of the NMR MLR they contain are all going to be different to that of the X-ray complex and MLR X-ray respectively. Since the energetic calculations are based in differences in structure, it is to be expected that differences in energetic parameters (*e.g.* G, H and S) are going to exist between the different MLR NMR and the MLR X-ray (ΔG°_{mlr}) and the different docked complexes and the X-ray complex (ΔG°_{comp}, see Figure 4).

We present in Figure 5 the calculated ΔG°_d of the docked complexes as a function of the positional r.m.s.d. Positional r.m.s.d. is defined as the r.m.s. difference between the location of the heavy atoms of the MLR in a docked complex as a function of their location in the X-ray complex when the PP-1c molecules are superimposed. Positional r.m.s.d. takes into account the structural and location differences of MLR between the complexes. One can clearly see that there is a general trend for the ΔG°_d to be larger as the r.m.s.d. becomes smaller. This relationship indicates that there is a clear tendency for the calculated ΔG°_d to be closer to the one calculated from the X-ray complex as the structure of the docked complexes become closer to it. There is, however, a significant scatter in the r.m.s.d., *i.e.* many complexes with higher r.m.s.d. have higher affinities than some with a lower r.m.s.d.. This is somewhat similar to the case of false positives encountered in docking experiment with potential energy force fields[11]. On the other hand, this could be an indication that there are, in fact, different possible configurations for the complexes that lead to similar decreases in free energy in solution. We also find that some of the docked complexes have a larger affinity that of the X-ray complex.

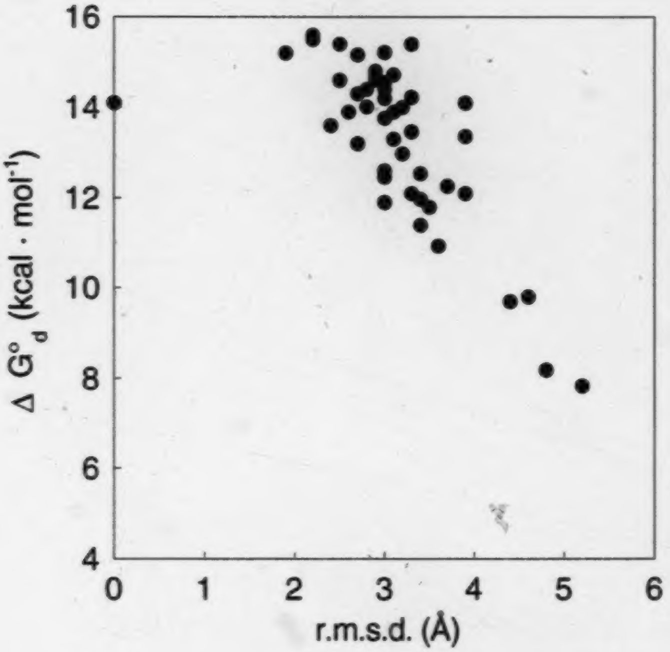


Figure 5- *The calculated $\Delta G°_d$ of the docked complexes as a function of the positional r.m.s.d. Positional r.m.s.d. is defined in the text.*

Formally, the docked complexes that give larger $\Delta G°_d$ than the X-ray complex must have lower free energy (more stable than the X-ray complex) and/or MLR NMR that have higher free energy (less stable) than the X-ray MLR. Since we make the assumption that the structure of PP-1c does not change upon dissociation, therefore its free energy is not changing either in the calculations. To explore the origins of the higher than X-ray complex $\Delta G°_d$, we calculated the $\Delta G°_{comp}$ for all the docked complexes and the $\Delta G°_{mlr}$ (see Fig. 4) for all the NMR MLR. Values for $\Delta G°_{comp}$ and $\Delta G°_{mlr}$ were calculated from the differences in ASA between the docked complexes and the X-ray complex and the NMR MLR and the X-ray MLR. A positive value for both $\Delta G°_{comp}$ and $\Delta G°_{mlr}$ indicate that the particular docked complex and NMR MLR are less stable that the X-ray complex and X-ray MLR respectively.

We present in Figure 6 the plot of the ΔG°_d as a function of ΔG°_{comp}. As one can observe, a clear correlation between ΔG°_d and ΔG°_{comp} exists *i.e.* the more stable the complex the more unfavorable the ΔG°_d. On the other hand, none of the complexes are more stable than the crystal structure (no negative ΔG°_{comp}). Therefore it is clear that the ΔG°_ds greater that that of the X-ray complex have to originate from MLR solution structures having positive ΔG°_{mlr}. The ΔG°_{mlr} were all found to be positive as expected from the argument above. This could to indicate that the X-ray MLR is a more stable conformation than all of the NMR MLR. Although this might be true, we notice that relatively small changes in ASA (*e.g.* $\Delta ASA_{pol} = 50$ Å^2 and $\Delta ASA_{np} = 50$ Å^2) lead to high ΔG°_{mlr} values (*e.g.* 2 kcal · mol^{-1}). These changes originate from small conformational changes (fluctuations) in the backbone and in side-chain dihedral angles that are unlikely to lead to such important changes in free energy as the one calculated (*e.g.* 2 kcal · mol^{-1}). These fluctuations should, to a first approximation, be nearly isoenergetic and not likely to affect significantly the relative population of the different members of the ensemble of solution structures. Unstable conformations of the toxin will increase the ΔG°_d but on the other hand will not reflect the most probable conformation of the toxin nor a realistic representation of the dissociation reaction. We assume here that the ensemble of NMR MLR structures are equally populated and consider that the ΔG°_{mlr} corresponds more or less to noise inherent to the present method to calculate differences in free energy between structurally fluctuating small peptides. We have therefore subtracted the corresponding ΔG°_{mlr} (and similarly differences in enthalpy ΔH°_{mlr} and entropy ΔS°_{mlr}) from the ΔG°_d (ΔH°_d and ΔS°_d) of the different docked complexes to yield a corrected ΔG°_d: $\Delta G^\circ_{d,corr}$ (Fig. 5).

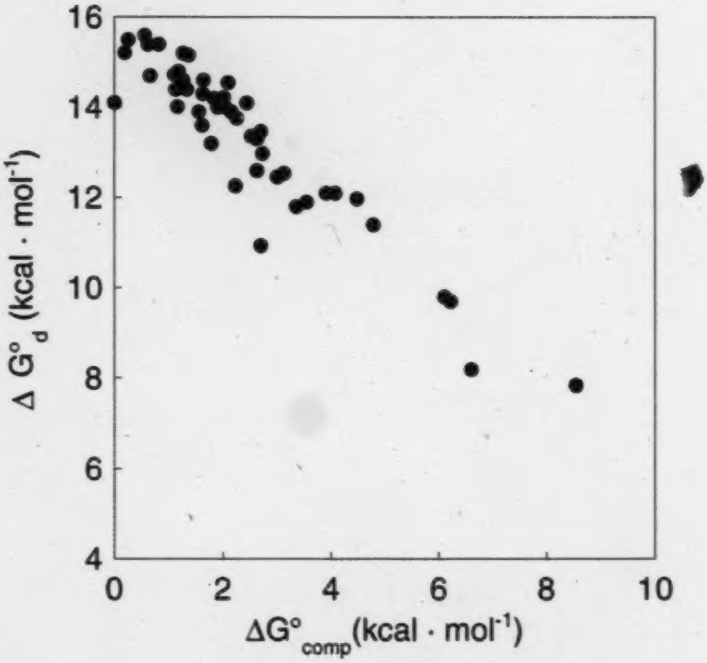


Figure 6- *A plot of the ΔG°$_d$ as a function of ΔG°$_{comp}$ of the complexes as described in Figure 4.*

We present on Figure 7 the plot of the corrected $\Delta G^\circ_{d.corr}$ as a function of ΔG°_{comp}. We can see that there is a nearly perfect correlation. This is explained from the fact that the relative $\Delta G^\circ_{d.corr}$ depends only on the relative ΔG°_{comp} or in other words the more stable the complex the larger the ΔG°_{d}.

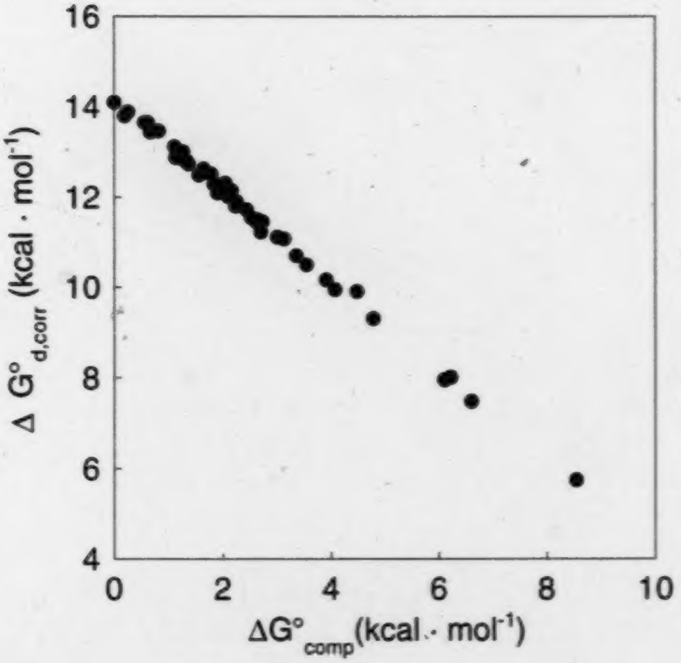


Figure 7- *The plot of the corrected $\Delta G^\circ_{d.corr}$ as a function of ΔG°_{comp}*

Finally, we present on Figure 8 the plot of $\Delta G^\circ_{d,corr}$ as a function of positional r.m.s.d. depicting that none of the $\Delta G^\circ_{d,corr}$ are larger than the ΔG°_d of the X-ray complex but still highlighting the scatter in r.m.s.d. This is particularly evident for the cluster (open circles) with $\Delta G^\circ_{d,corr}$ closer to the ΔG°_d of the X-ray complex.

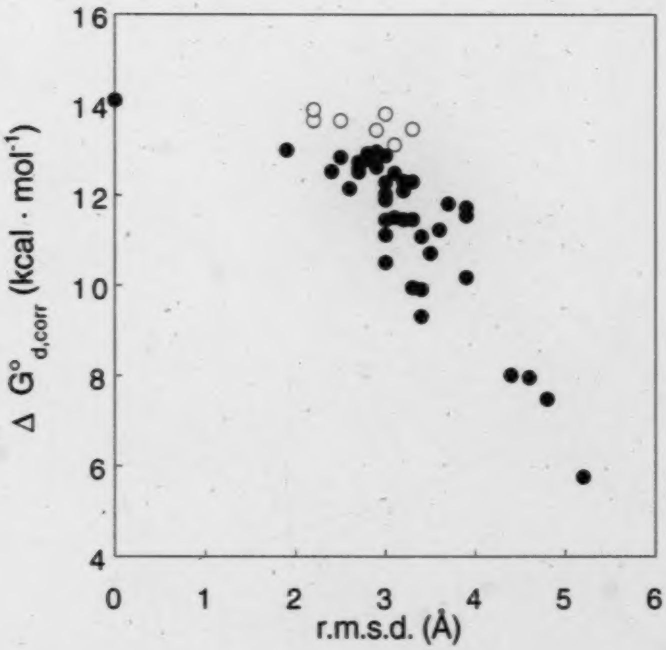


Figure 8- The plot of the corrected $\Delta G^\circ_{d,corr}$ as a function of positional r.m.s.d. See text for further details.

Enthalpy-entropy compensation phenomenon evidenced from structure-based thermodynamics

Figure 9 displays the structures of the 6 complexes highlighted in Figure 8 as open circles including the average structure (purple). The X-ray MLR is also displayed (yellow). One can notice that the backbones of all the MLR molecules are close to each other and that the high r.m.s.d. values come from differences in the conformations of the long Adda and Arg side-chains. With this regard, one can see that NMR MLR have Adda side-chains that are more exposed and Arg side-chains that lie closer to the enzyme potentially forming H-bonds and/or salt bridges with either Glu 275 and/or Asp 220 (Fig. 2). It is clear the present parameterization is able to recognize docked complexes that are close to the X-ray complex on the basis of the $\Delta G°_{d.coor}$ calculated here but it still show some significant structural differences. This could indicate that this approach does not have a high degree of discrimination. On the other hand, it is possible that the complexes with $\Delta G°_{d.coor}$ similar to the X-ray complex could be populated in solution and that the differences in structure displayed in Figure 9 are nearly isoenergetic.

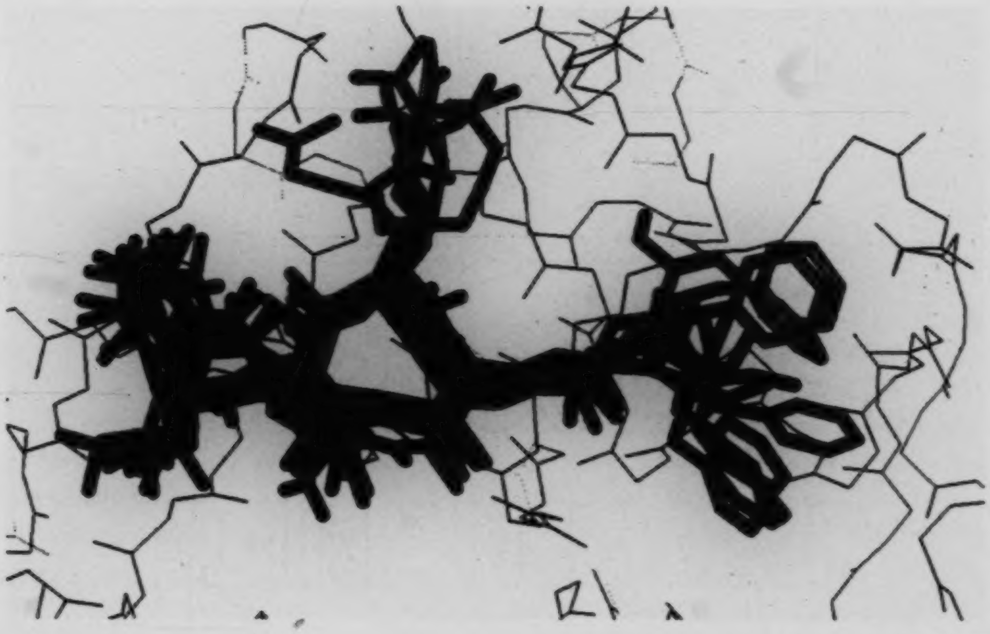


Figure 9- *The structure of the seven complexes representing the highest $\Delta G_{d.corr}$ highlighted in Figure 8 including the bound crystal structure of MLR (white) and the geometric average of the 46 NMR MLR (purple). The X-ray PP-1c molecules were superimpose onto one another.*

In order to investigate why complexes with relatively high and scattered r.m.s.d. can give rise to similar $\Delta G°_{d}$ (similar free energies), we have analyzed the normalized $\Delta H°_{d.corr}$ and $\Delta S°_{d.corr}$ (corrected for the variations in the enthalpy and entropy of the different NMR MLR) and the structural features of the complexes of the high $\Delta G°_{d.corr}$ cluster. We present on Figure 10A a plot of the $\Delta H°_{d.corr}$ as a function of the $\Delta S°_{d.corr}$ and on Figure 10B a plot of $\Delta H°_{comp}$ as a function of the $\Delta S°_{comp}$ calculated for the subset of seven complexes. The larger the r.m.s.d. of the complexes, the more they lie on the left right of the plots. One can clearly observe a linear relationship between the two quantities in both instances. This indicates the existence of an enthalpy-entropy compensation phenomenon in the calculations. Experimental enthalpy-entropy compensation effects are ubiquitous and have been reported for the binding of series of ligands to their specific enzymes or for binding studies carried on at different pHs or ionic strengths[47,48]. Experimental enthalpy-entropy compensation have also been reported for the stability of protein mutants[49,50]. It is thought that water plays a key role in the mechanism of the compensation effect with typical compensation temperatures or T_c (slope of the ΔH°_d vs ΔS°_d curves) between 270 and 320 K[47,48]. Eftink et al.[51] have presented a thermodynamic model for the enthalpy-entropy compensation in ligand-enzyme systems. In their model the compensation can be modeled if ligands can bind different microstates of the enzyme with different affinities that can vary under the different experimental conditions. Such a model is also applicable to different microstates of the ligand[51]. The compensation effect noted here is somewhat different. In fact, we observe that different 'microstates' of a ligand can have the same affinity for one conformation (or microstate) of an enzyme and that different 'microstates' of a complex can have the same stability under one 'experimental' condition. The fact that the calculations give rise to the compensation should inform us on the mechanism of the experimental compensations observed as well as on the relevance of our calculations in the understanding of the dissociation of protein complexes in solution.

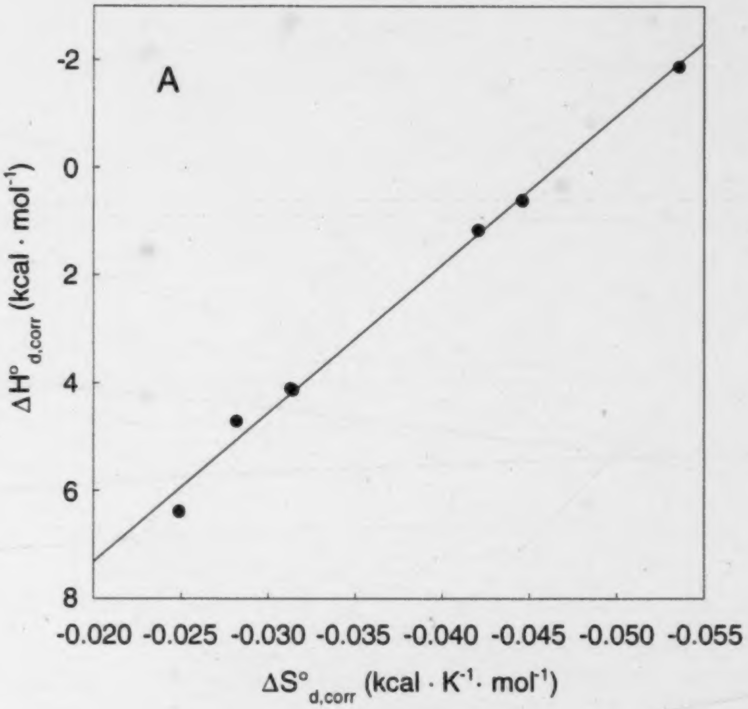


Figure 10A- *Plot of the $\Delta H_{d,corr}$ as a function of the $\Delta S_{d,corr}$ for the seven complexes represented in Figure 9. The slope of the line is 280K.*

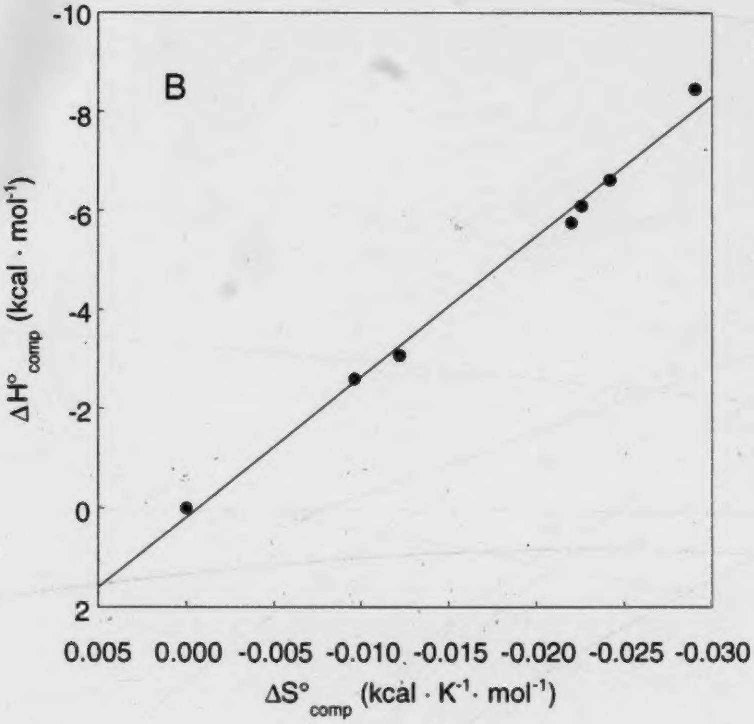


Figure 10B- *Plot of the ΔH_{comp} as a function of the ΔS_{comp} for the seven complexes presented in Figure 9. The slope of the line is 281K.*

Since the main structural differences are located at the Arg and Adda side-chains (Fig. 9), we will focus on the implications of these side-chains in a putative compensation mechanism. For the dissociation of docked complexes with the Adda side-chain more solvent exposed and the Arg side-chain lying closer to PP-1ç (forming salt-brides and/or H-Bonds, see Fig. 2) compared to the X-ray complex, the relative ratio $\Delta ASA_{pol}/\Delta ASA_{np}$ will increase. This leads to a more unfavorable ΔH°_d(25). In other words, the enthalpy of such a docked complex (PP-1c:MLR and the solvent) will decrease compared to the X-ray complex. Indeed, the proportion of polar and more enthalpically favorable[52] interactions per Å^2 of buried surface in the complex will increase. Therefore more positive (unfavorable) $\Delta H°_d$ for the docked complexes are obtained. On the other hand, we observe that the reduction in entropy of the solvent upon dissociation, $\Delta S°_{solv}(25)$, becomes less unfavorable for such complexes. In other words the entropy of the complexes (PP-1c, MLR and the solvent) will decrease. Variations in ΔS_{conf} for the $\Delta G°_{d,corr}$ and $\Delta G°_{comp}$ were found to be about one order of magnitude lower than that of the ΔS_{solv}. Therefore in the system studied here (PP-1c, MLR and the solvent), the response to a decrease in enthalpy (more favorable PP-1c:MLR interactions) is a decrease in entropy of water. The variations in enthalpy ($\Delta H°_d(25)$ or $\Delta H°_{comp}(25)$) are almost exactly compensated by the variations in entropies ($-T \cdot \Delta S°_d(25)$ or $-T \cdot \Delta S°_{comp}(25)$) giving rise to the linear relationships observed (Fig. 10) and leaving $\Delta G°_{d,corr}$ or $\Delta G°_{comp}$ unchanged. It is interesting to note that the slopes of Figure 7A and B are equal to approx. 280 K, which is of the same order of magnitude of the values for T_c values reported in the literature and which were attributed to the implication of water[47,48]. The coincidence in T_c values could be an indication that the compensation mechanism described above is realistic as far as the role of water and that the parameterization used here is faithful enough to simulate entropy-enthalpy compensations that occur in the PP-1c:MLR complex in solution (and complex dissociation). The mechanism described could also apply for the compensation effects observed under different experimental conditions referred to above. The existence of structural and energy fluctuations in proteins are both documented on theoretical[53] and experimental backgrounds[54]. These fluctuations remind us that the structure and thermodynamic parameters (such as the enthalpy of dissociation of a complex) observed for proteins and protein complexes are weighted mean values or ensemble averages.

Cooperativity of dissociation and multiple conformations (microstates) for the PP-1c:MLR complex

According to Boltzmann's statistics the docked complexes that give rise to the largest $\Delta G°_d$ (or that are the most stable) should correspond to the most probable conformations within the limited ensemble of complexes generated here. We have calculated the relative population (probabilities) of the complexes according to the following equation:

$$P_i = \frac{\exp(-\Delta G_{comp,i} / RT)}{\sum_i \exp(-\Delta G_{comp,i} / RT)}, \qquad (11)$$

where $\Delta G^{\circ}_{comp,i}$ is the relative free energy of the complex compared to the crystal structure, R is the gas constant and T is the absolute temperature.

We present in Figure 11A the P_i of every complex as a function of the positional r.m.s.d. of all heavy atoms of the MLR. As can be seen there is an increase in P_i as the r.m.s.d. reaches around 3 Å. But one can also notice the scatter in r.m.s.d. As discussed before, this is a manifestation of the enthalpy-entropy compensation. We show in Figure 11B the same P_i as a function of the positional r.m.s.d. of the backbone of the toxin. We can see that the increase is sharper indicating a cooperative role for the cyclic portion of the MLR molecule. This is quite interesting because the cyclic portion (Masp, D-Glu) is involved in the molecular recognition. Since the D-Glu and Masp side-chains have only one rotable bond (Fig. 1 top), the position of the backbone controls the position and the burial of the two carboxylic groups at the interface of the complex. Therefore Figure 11B demonstrates that if the residues that are involved in making the specific interactions are not properly buried by not being in the right position, the probability of the particular complex will be low. Moreover Figure 11A indicates that the structure of hydrophobic side-chains (*e.g.* Adda) are more likely to fluctuate or to have different conformations in the significantly populated complexes. Therefore Figure 9 could represents a dynamic rendition of the complex on a short time scale as this type of motion occurs between states of similar free energy. It has to be noticed that this rendition would be partial as all the different possible configurations are not present in the limited ensemble of complexes used here. The amplitude of the motions seen on Figure 9 would have to be reconciled with the more 'static' structure of MLR in the X-ray complex but this effort is beyond the scope of the present study.

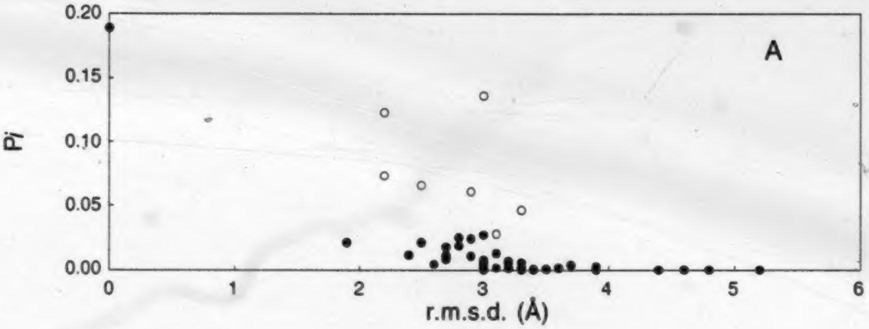


Figure 11A- *The P_i of every complex as a function of the positional r.m.s.d. of all heavy atoms of the MLR. P_i is defined in the text.*

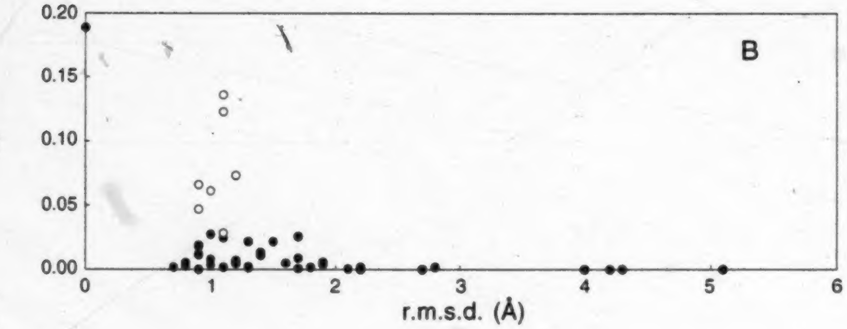


Figure 11B- The same probabilities as in A. except this time as a function of the positional r.m.s.d of the backbone of the toxin. The open circle refer to the complexes on Fig. 9.

In order to confirm our results that the complex between PP-1c and MLR shows significant structural fluctuations and characterize their amplitude, NMR relaxation studies should be done on MLR bound to PP-1c. Experimental efforts towards such measurements are being pursued in our laboratory but are being complicated by the high molecular weight of the complex and solubility issues. Experimental thermodynamic measurements could also shed light on the existence and the extent of the structural fluctuations. Indeed, if the experimental ΔH°_{d} for the PP-1c:MLR complex (corrected for proton transfer effects) could be obtained and observed to be more unfavorable than the one calculated from the X-ray complex this would indicate the population of different microstates of the complexes with different enthalpies. Recalling that the ΔH°_d measured would correspond to a weighted mean value or ensemble average ($\langle \Delta H^\circ_d \rangle$):

$$\langle \Delta H^o_{d,corr} \rangle = \sum_i P_i \cdot \Delta H^o_{d,corr,i} \tag{12}$$

where $\Delta H^\circ_{d,corr\,i}$ is the enthalpy of dissociation for each complex or microstates of an ensemble. For example the $\langle \Delta H^\circ_{d,corr} \rangle$ for the limited ensemble of complexes generated here is 2.05 kcal · mol^{-1} compared to –1.87 kcal · mol^{-1} calculated from the X-ray complex.

Implications for molecular recognition

It is clear from our results that different conformations (microstates) than the X-ray MLR can bind to one microstate of PP-1c (X-ray PP-1c) with similar affinities. The major structural differences between these microstates are located at side-chains (*e.g.* Adda and Arg) that are not involved in forming specific interactions as evidence from the X-ray complex (Fig. 2). On the other hand the side-chains that are involved in molecular recognition (Masp and D-Glu) occur in the best defined region of MLR in solution (Fig. 1). The fact that the portion or domain responsible for molecular recognition is already folded is an advantage for effective rates of binding (k_{on}). Indeed, all the microstates are potentially able to recognize PP-1c. In addition, the fact that the Adda and Arg side-chains can exist many different conformations without affecting the ΔG°_d should also help for elevated k_{on} rates. There is a limit, however, to which the Adda side-chain can vary its conformation without changing ΔG°_d significantly. Indeed, a complex generated by manually changing the conformation of the Adda side-chain in the X-ray complex so that it is maximally exposed had a ΔG°_d decreased by 2 kcal · mol^{-1} or a 100 fold decreased K_{Da} (data not shown).

Our results might also have general implications for the study of free ligands by solution state NMR. Indeed, our study indicates that all the members of the ensemble of structures usually calculated should be considered as being potentially able to bind a receptor and not only the geometric average. Although the geometric average structure of MLR here was found to be in the cluster of 7 complexes of high ΔG°_d, it could well not have been. In fact, for the calculation of a geometric average all the members of the ensemble are given the same weight and therefore can lead to a biased structure by conformers that may not be populated in solution although they satisfy the experimental

NMR restraints. The best defined regions, as evidenced from the ensemble, could potentially be involved in molecular recognition. Finally, and as shown here, the worst defined regions of a ligand can contribute significantly to the affinity of binding and as such can not be ignored for their potential importance in the binding process.

In conclusion, using a structure-based approach we have calculated from the X-ray complex of PP-1c and MLR a free energy of dissociation that is in close agreement with the reported K_i. We also note that the residues with the largest contribution to the overall dissociation free energy to be Arg-96, Arg-221 on PP-1c and D-Glu and Adda on MLR. This is in accordance with experimental data that shows that they are indeed critical for complex formation. We notice that the reported high affinity of MLR is due to its cyclic nature, *i.e.* the dissociation reaction is not linked to a conformational change in the backbone and therefore leads to a minimal conformational entropy gain compared to a fictitious linear version of the toxin. The analysis of 47 complexes obtained from docking a family of NMR solution structures of MLR onto the crystal structure of the PP-1c, predicts structural fluctuations for the bound form of MLR especially for the long Adda and Arg side-chains. This analysis also suggests that the cyclic part of the MLR is more important that the long hydrophobic Adda side-chain in the cooperativity of dissociation (binding). We observe an enthalpy-entropy compensation phenomenon for which we describe a putative mechanism that could be applicable for experimentally observed compensation effects and to understand the mechanisms of structural fluctuations and molecular dynamics in solution.

Analysis of NMR docked MLR:PP-1c complexes: Part 1-All complexes

Figure 12 is the free energy of dissociation for every complex including crystal (C). The average free energy of dissociation is 13.3 kcal/mol, with the standard deviation being 1.8 kcal/mol and standard error being 0.3 kcal/mol. There are 16 complexes that are significantly below 13.3 kcal/mol, are therefore relative to the rest of the complexes would have a significantly reduced affinity for binding to PP-1c. The remaining 32 complexes have average or higher (13 kcal/mol) free energy of dissociations indicating that these are of higher affinity complexes.

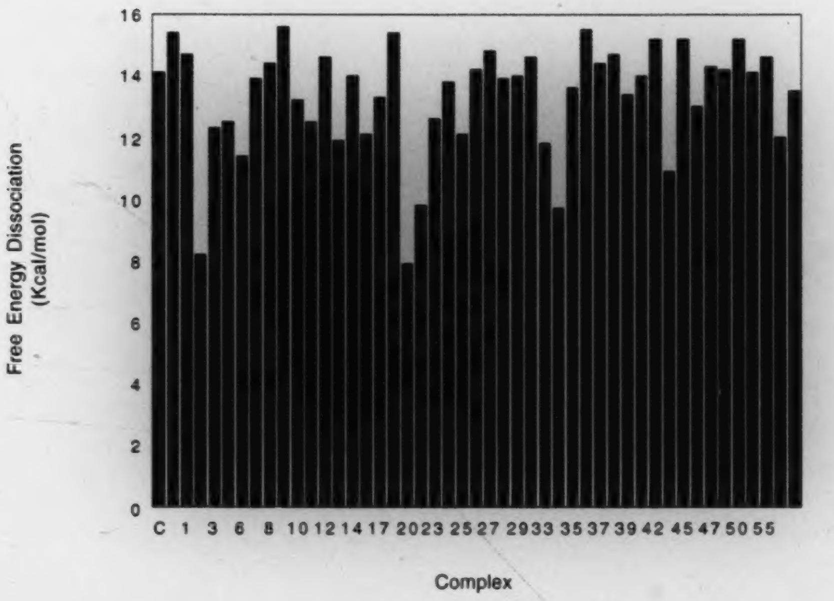


Figure 12- *Free energy of dissociation for every docked NMR MLR:PP-1c complex as well as the crystal structure complex. C is the crystal structure complex (as with all other figures). The average free energy of dissociation is 13.3 kcal/mol, with the standard deviation being 1.8 kcal/mol and standard error being 0.3 kcal/mol.*

We present in Figure 13 (*left*) 30 calculated NMR structures, the average solution structure and the crystal structure of MLR docked or bound to PP-1c. These complexes have average or better than average free energy of dissociations (Figure 12) and are classified as being high affinity complexes. From this group of 30 calculated structures and additionally the average solution structure and the bound crystal structure, Figure 13 (*right*) displays the structures of the seven complexes representing the highest ΔG_d or ΔG_{corr} cluster highlighted in Figure 8. One can notice that the backbones of all the MLR molecules are close to each other and that the high r.m.s.d. values come from differences in the conformations of the long Adda and Arg side-chains. With this regard, one can see that some of the MLR have Adda side-chains that are more exposed and Arg side chains that lie closer to the enzyme. The MLR structure in two of lowest binding free energy complexes (purple) have overall structures that are very similar to that of the of the crystal structure (yellow) except for the Arg side-chain. It is clear the present parameterization is able to recognize complexes that are close to the target on the basis of empirical free energy of dissociation but still show some significant difference.

Figure 13 (left)- The 30 calculated NMR structures, the average solution structure and the crystal structure of MLR docked or bound to PP-1c. These complexes have ΔG_d greater than or equal to 13 kcal/mol and are categorized as being the high affinity complexes. (right)- The structures of the seven complexes representing the highest ΔG_d or ΔG_{corr} cluster highlighted in Figure 8 excluding the lowest open circle.

Figure 14 is the corrected free energy of dissociation for all the complexes including the crystal complex. This takes into account the variability in free energy of the free ligand (See Figure 4). The average corrected free energy of dissociation is 11.7 kcal/mol, with the standard deviation and standard error being 1.8 kcal/mol and 0.3 kcal/mol respectively. As seen in the figure, the crystal complex has the highest corrected free energy of dissociation. As in Figure 12, the 16 lowest free energy complexes have also the lowest corrected free energies and with such low affinity would probably not be realistically able to bind.

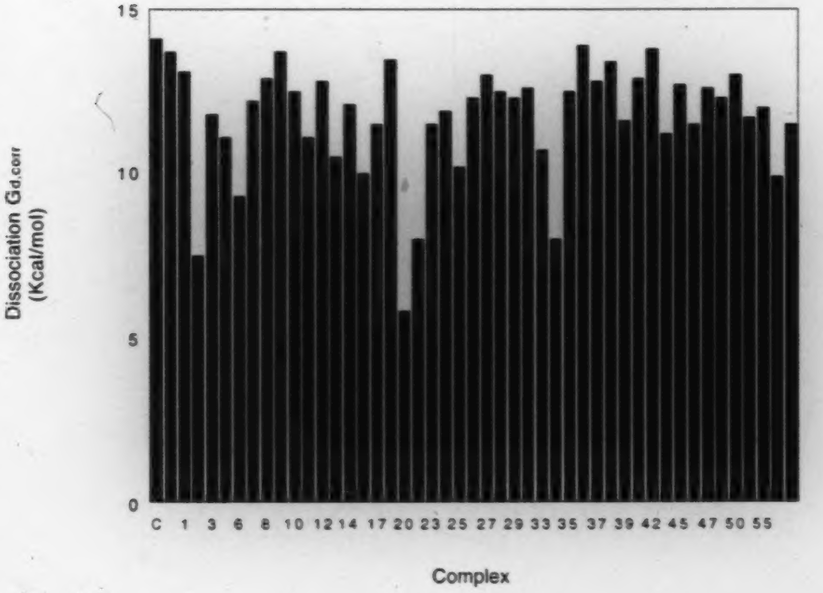


Figure 14- *Corrected dissociation free energy for every docked NMR complex and crystal complex. The correction is based on removing the free energy differences in the free ligands (Figure 4). The average corrected free energy of dissociation is 11.7 kcal/mol, with the standard deviation and standard error being 1.8 kcal/mol and 0.3 kcal/mol respectively.*

Figure 15 is the entropy and enthalpy contributions to the dissociation free energy for all the complexes. The average dissociation entropy is 10.9 kcal/K/mol (standard deviation is 2.0 kcal/K/mol, standard error is 0.3 kcal/K/mol). The average dissociation enthalpy is 2.4 kcal/mol (standard deviation is 2.4 kcal/mol, standard error is 0.3 kcal/mol). Comparing the averages and taking into account standard deviations, it can be concluded that for the vast majority of complexes the binding is being driven mostly by entropy. Enthalpy only has a minor favorable impact on microcystin-LR binding in the majority of complexes. Some complexes, such as the crystal complex, have negative dissociation enthalpies. This implies that those complexes would have unfavorable binding enthalpies, where entropy would be required to overcome the enthalpy in order for binding to take place. Another notable feature of this figure is the entropy-enthalpy compensation effect. Most of the complexes have a free energy of dissociation between 13-16 kcal/mol. However, there is a wide variety of entropies and enthalpies that add up to result in the different free energies. A better example will be seen in Figure 20.

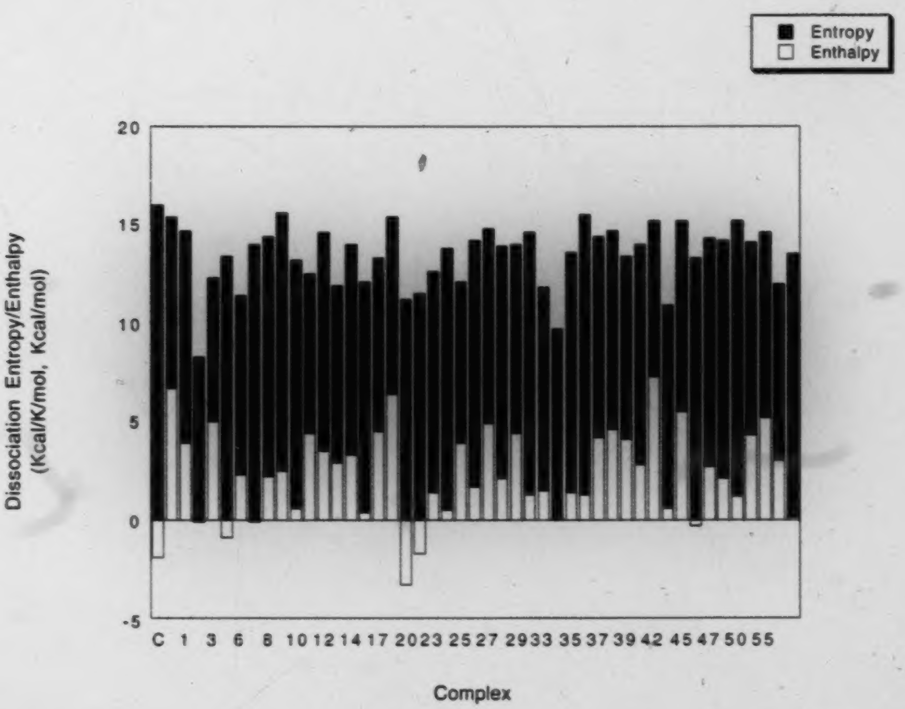


***Figure 15-** Dissociation entropy and enthalpy for every complex. The average dissociation entropy is 10.9 kcal/K/mol (standard deviation is 2.0 kcal/K/mol, standard error is 0.3 kcal/K/mol). The average dissociation enthalpy is 2.4 kcal/mol (standard deviation is 2.4 kcal/mol, standard error is 0.3 kcal/mol).*

Figure 16 is the burial of accessible surface area by microcystin-LR for every complex including crystal. The average nonpolar burial of accessible surface area by microcystin-LR is 403 Å^2, with the standard deviation being 52 Å^2 and standard error being 8 Å^2. The average polar burial of accessible surface area by microcystin-LR is 169 Å^2, with the standard deviation being 33 Å^2 and the standard error is 5 Å^2. Microcystin-LR buries in most cases over two times the amount of nonpolar area in comparison to polar area. This definitely relates to why the binding of microcystin-LR to PP-1c is mostly entropically driven as observed in Figure 15. The variability in nonpolar and polar surface area buried from one complex to the next is interesting to note. A more detailed investigation of this variability will be discussed later.

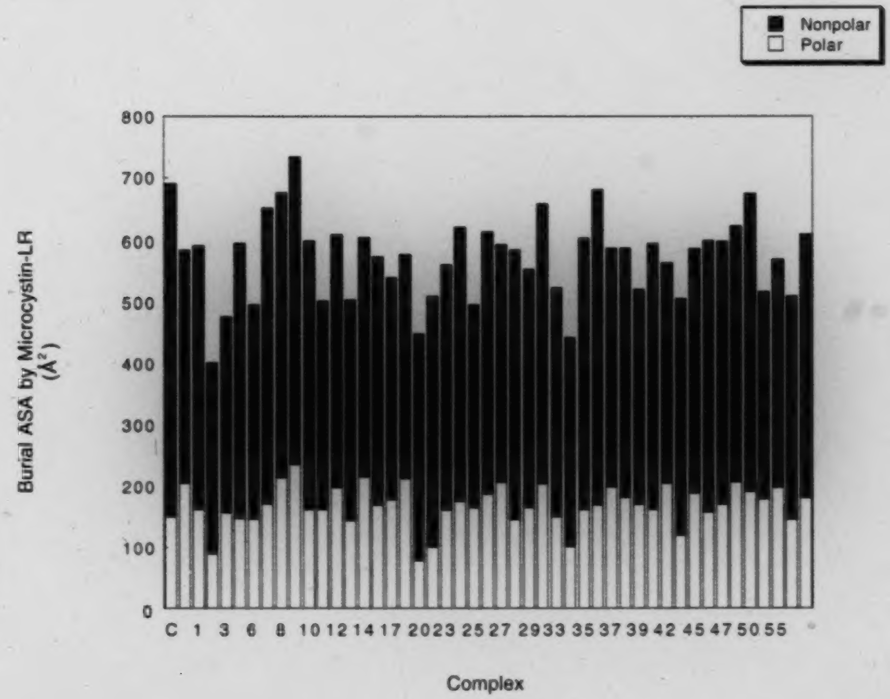


Figure 16- *Burial of accessible surface area by microcystin-LR in every complex. In most complexes, there is over twice as much burial of nonpolar area verses polar area. The average nonpolar burial of accessible surface area by microcystin-LR is 403 Å^2, with the standard deviation being 52 Å^2 and standard error being 8 Å^2. The average polar burial of accessible surface area by microcystin-LR is 169 Å^2, with the standard deviation being 33 Å^2 and the standard error is 5 Å^2.*

Figure 17 shows the burial of accessible surface area of PP-1c upon microcystin-LR attachment for every complex. The average nonpolar accessible surface area burial of PP-1c is 245 Å^2 (standard deviation = 32 Å^2, standard error = 5 Å^2). The average polar accessible surface area burial of PP-1c is 277 Å^2 (standard deviation = 31 Å^2, standard error = 5 Å^2). Overall, there is slightly less total accessible surface area burial of PP-1c in comparison to microcystin-LR. In contrast to microcystin-LR, PP-1c buries approximately the same amount of nonpolar and polar surface area (averages within 1 standard deviation of each other). Additionally, there a reduction in the amount of variability in the nonpolar surface area buried when compared to microcystin-LR which shows a higher than normal standard deviation. Therefore, PP-1c is more neutral in its contribution to binding enthalpy and entropy upon being inhibited by microcystin-LR.

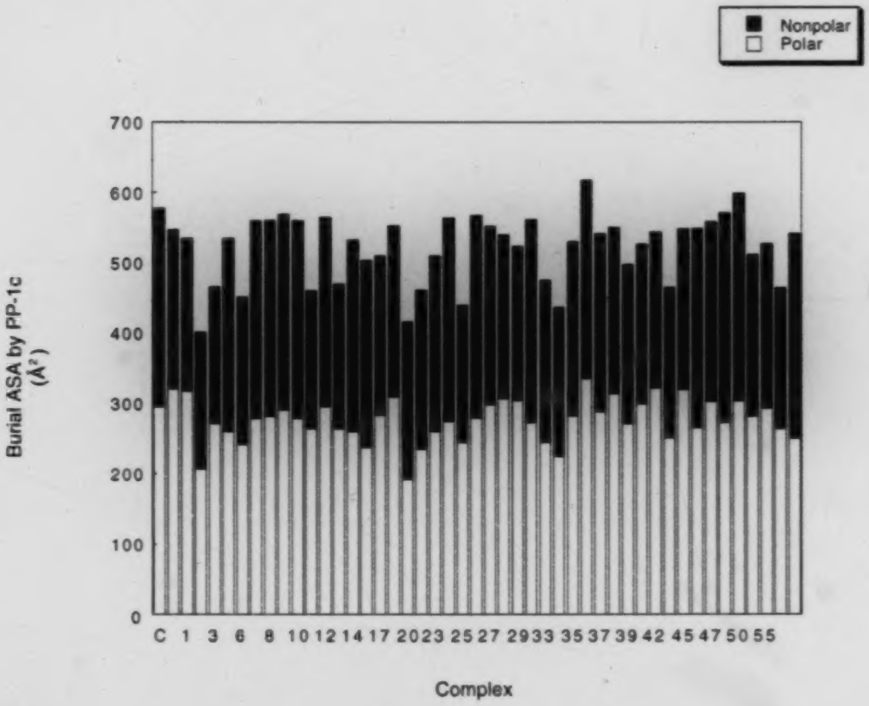


Figure 17- *Burial of accessible surface area of PP-1c upon microcystin-LR attachment for every complex. The average nonpolar accessible surface area burial of PP-1c is 245 Å^2 (standard deviation = 32 Å^2, standard error = 5 Å^2). The average polar accessible surface area burial of PP-1c is 277 Å^2 (standard deviation = 31 Å^2, standard error = 5 Å^2). There is roughly equivalent burial of nonpolar verses polar area, and less variability in the amount of nonpolar area buried in comparison to microcystin-LR (see Figure 16)*

Analysis of NMR docked MLR:PP-1c complexes:Part 2-32 High Affinity Complexes

Out of the 47 docked NMR complexes (including NMR average), 16 complexes had significantly less than average free energies of dissociation (Figure 12). The remaining 30 complexes, NMR average complex, and crystal complex had the significantly highest free energies of dissociation (Figure 13 *left*). This section will focus in more detail on the attributes of these highest dissociation free energy complexes.

Figure 18 shows the free energy of dissociation for the 30 high affinity complexes, average NMR complex, and crystal structure complex. The average free energy of dissociation for this set of complexes is 14.3 kcal/mol (standard deviation = 0.7 kcal/mol, standard error = 0.1 kcal/mol). There is less standard deviation in this set of data in comparison to the complete set (Figure 12). Based on standard error, there is a significant increase in the average free energy of dissociation verses the complete set of data.

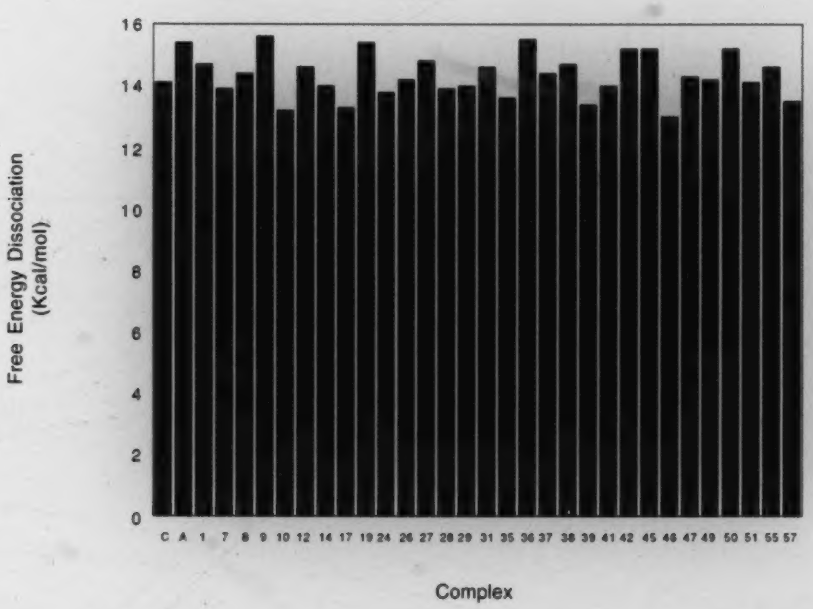


Figure 18- *Free energy of dissociation for the 30 best complexes, average NMR complex, and crystal structure complex. The average free energy of dissociation for this set of complexes is 14.3 kcal/mol (standard deviation = 0.7 kcal/mol, standard error = 0.1 kcal/mol).*

Figure 19 is the corrected free energy of dissociation for the 30 high affinity calculated NMR docked complexes, average NMR complex, and crystal complex. The average corrected free energy of dissociation for this set of complexes is 12.7 kcal/mol (standard deviation = 0.7 kcal/mol, standard error = 0.1 kcal/mol). As with Figure 14, the crystal structure has the highest corrected free energy of dissociation once the free energy of dissociation for the free ligand is taken into account.

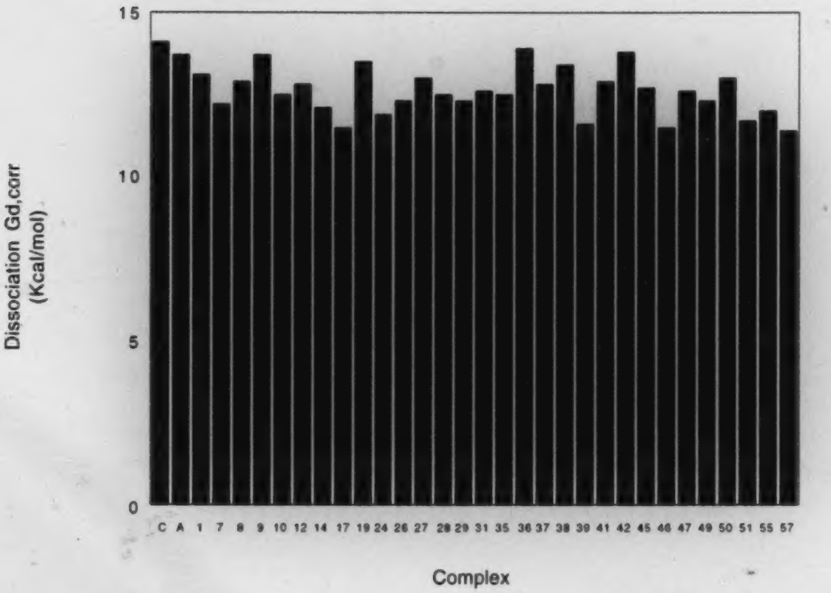


Figure 19- *Corrected free energy of dissociation for the 30 high affinity calculated NMR docked complexes, average NMR complex, and crystal complex. The average corrected free energy of dissociation for this set of complexes is 12.7 kcal/mol (standard deviation = 0.7 kcal/mol, standard error = 0.1 kcal/mol).*

Figure 20 is an excellent description of the entropy-enthalpy compensation also shown in Figure 10 and Figure 15. This figure shows the dissociation entropy and dissociation enthalpy for the 30 high affinity complexes plus average and crystal complexes. The average dissociation entropy is 11.4 kcal/K/mol, (standard deviation = 2.0 kcal/K/mol, standard error = 0.3 kcal/K/mol), while the average dissociation enthalpy is 2.9 kcal/mol (standard deviation = 2.2 kcal/mol, standard error = 0.4 kcal/mol). There is a clear dominance of favorable binding entropy when MLR inhibits PP-1c. However, while the free energy of dissociation is relatively stable in this set of complexes, the individual contributions from entropy and enthalpy are highly variable. Therefore, there are multiple possible combinations of entropy and enthalpy that can result in tight binding of MLR to PP-1c. Note that these changes in entropy and enthalpy are the result of only minor changes in the saddle-shaped backbone of MLR, and fluctuations in the side chain orientation of Adda and Arg once bound. More detail will be given on individual residues later in this section.

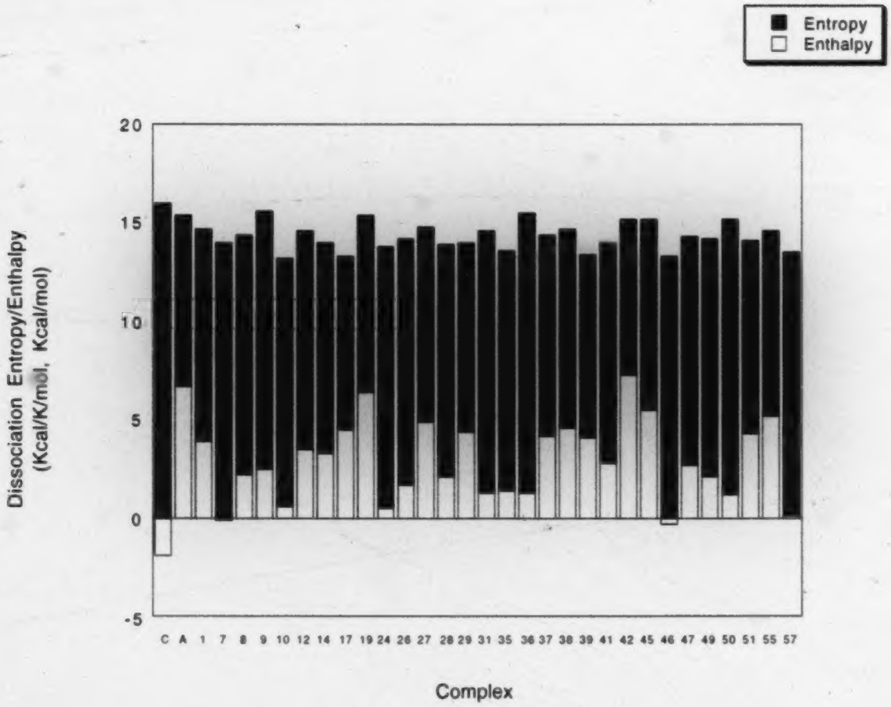


Figure 20- *The dissociation entropy and dissociation enthalpy for the 30 high affinity complexes plus average and crystal complexes. The average dissociation entropy is 11.4 kcal/K/mol, (standard deviation = 2.0 kcal/K/mol, standard error = 0.3 kcal/K/mol), while the average dissociation enthalpy is 2.9 kcal/mol (standard deviation = 2.2 kcal/mol, standard error = 0.4 kcal/mol).*

Is there any consistent pattern in the binding of MLR to PP-1c in these high dissociation energy complexes? Figure 21 answers this question by displaying the burial of accessible surface area of MLR upon complex formation for these complexes. The average burial of nonpolar area is 420 Å^2 (standard deviation = 49 Å^2, standard error = 9 Å^2), while the average burial of polar area is 185 Å^2 (standard deviation = 22 Å^2, standard error is 4 Å^2). There is variation in the burial of nonpolar and polar accessible surface area, indicating that MLR is able to tightly bind in a variety of similar structures/positions despite differing nonpolar and polar surface interactions. Overall, MLR is the major contributer to nonpolar surface area buried which is the reason why the binding is entropically driven.

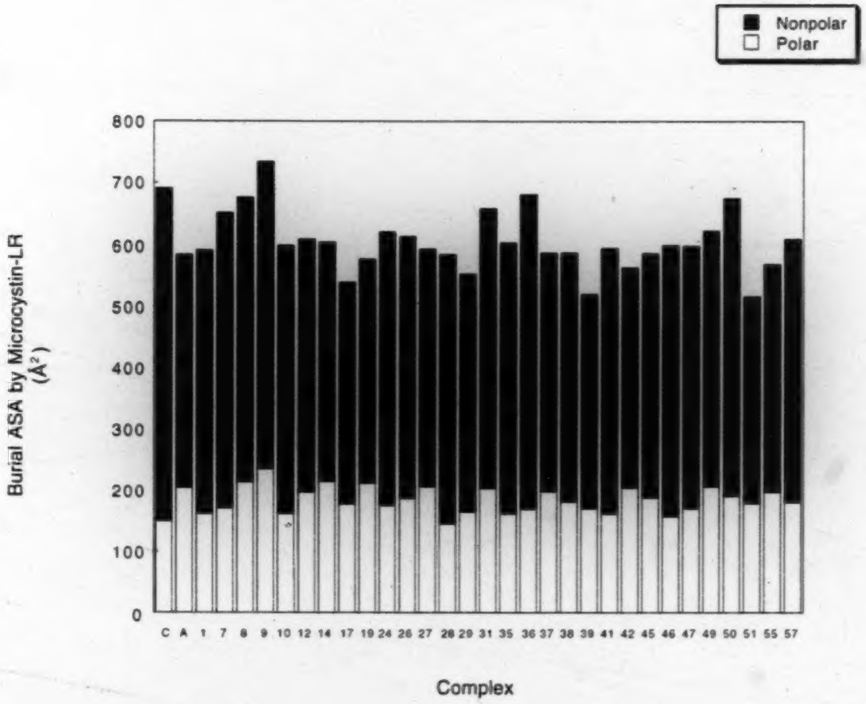


Figure 21- *Burial of accessible surface area of MLR upon complex formation for the 32 tightest binding complexes. The average burial of nonpolar area is 420 Å^2 (standard deviation = 49 Å^2, standard error = 9 Å^2), while the average burial of polar area is 185 Å^2 (standard deviation = 22 Å^2, standard error is 4 Å^2).*

Similar to MLR, PP-1c shows a variety of accessible surface area burial patterns when inhibited by MLR in the 30 best complexes (Figure 22). The average burial of nonpolar area by PP-1c is 257 Å^2 (standard deviation = 27 Å^2, standard error = 5 Å^2), while the average burial of polar area is 292 Å^2 (standard deviation = 20 Å^2, standard error = 4 Å^2). More polar area is generally buried than nonpolar for PP-1c, but they are in close proportions. There tends to be less variability in the nonpolar area buried in PP-1c than MLR.

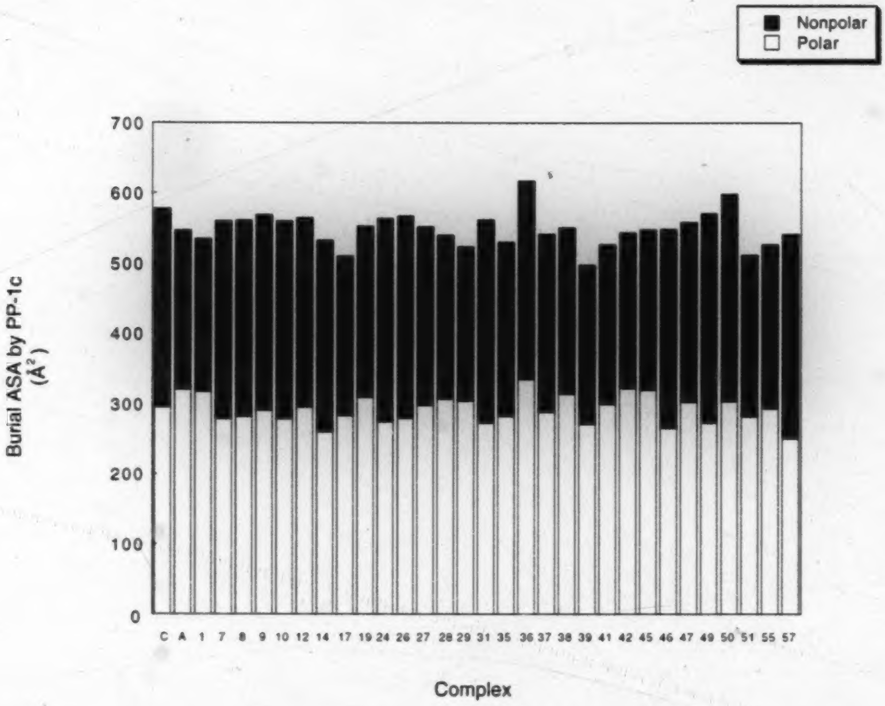


Figure 22- *Burial of accessible surface area by PP-1c in the 30 best calculated NMR complexes. The average burial of nonpolar area by PP-1c is 257 Å^2 (standard deviation = 27 Å^2, standard error = 5 Å^2), while the average burial of polar area is 292 Å^2 (standard deviation = 20 Å^2, standard error = 4 Å^2).*

Figure 23 is a close view of the Adda residue accessible surface area burial of microcystin-LR for the 30 best NMR calculated docked complexes. The average nonpolar burial of accessible surface area by Adda is 216 Å^2 (standard deviation = 47Å^2, standard error = 25Å^2), while the polar burial by Adda is only 25 Å^2 (standard deviation = 3 Å^2, standard error = 1 Å^2). When comparing the Adda nonpolar burial with the overall MLR nonpolar burial it is clear that the Adda residue is responsible for over half the total burial of nonpolar accessible surface area by MLR. However, there is a rather large variation in Adda nonpolar burial in the complexes observed in Figure 23. This can be observed structurally in the variety of different Adda side chain conformations in Figure 9 (*left*). Despite this variation all of these complexes have similar binding affinities. One can observe a correlation between the amount of nonpolar area buried by Adda and the degree of entropic contribution to binding. Thus, those complexes having a lesser degree of nonpolar Adda accessible surface area buried have enthalpy compensating with a stronger enthalpic contribution towards binding.

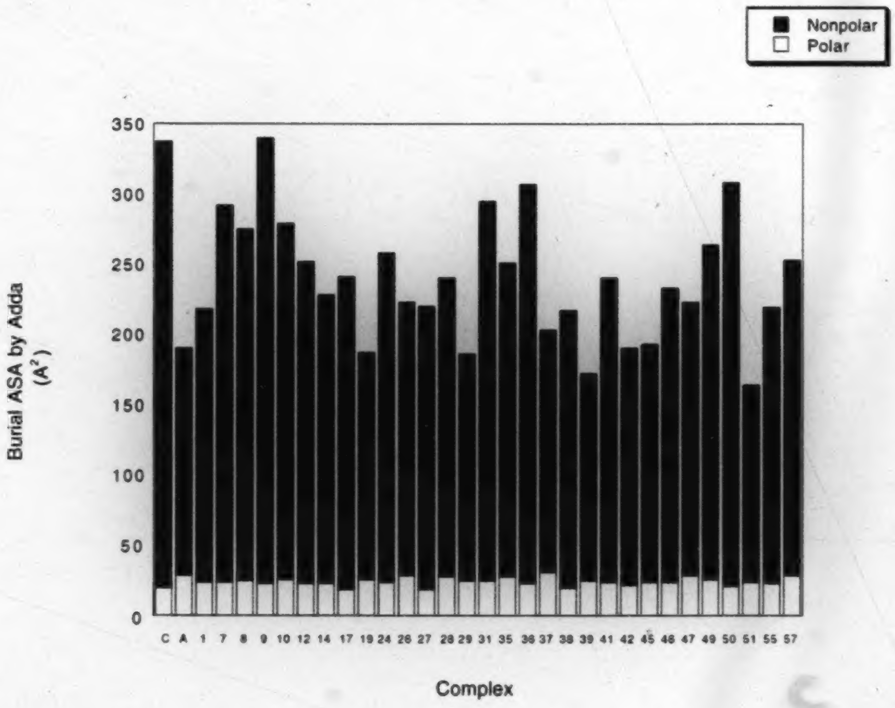


***Figure 23-** MLR's Adda residue accessible surface area burial for the 30 best NMR calculated docked complexes. The average nonpolar burial of accessible surface area by Adda is 216 Å^2 (standard deviation = 47Å^2, standard error = 25Å^2), while the polar burial by Adda is only 25 Å^2 (standard deviation = 3 Å^2, standard error = 1Å^2).*

Opposite to the strongly hydrophobic residue Adda in MLR is the hydrophilic residue Arg. Figure 24 displays the accessible surface area buried by the Arg residue of MLR in the 30 best calculated NMR complexes. The average nonpolar burial by Arg is 29 Å^2 (standard deviation = 12 Å^2, standard error = 2 Å^2), while the average polar area is 48 Å^2 (standard deviation =20 Å^2, standard error = 4 Å^2). The radical variation in the amount of nonpolar and polar area buried for the Arg residue is in tune with the large variety of structural conformations this residue has when bound (Figure 13). Despite the high standard deviations the complexes all have similar free energies of binding.

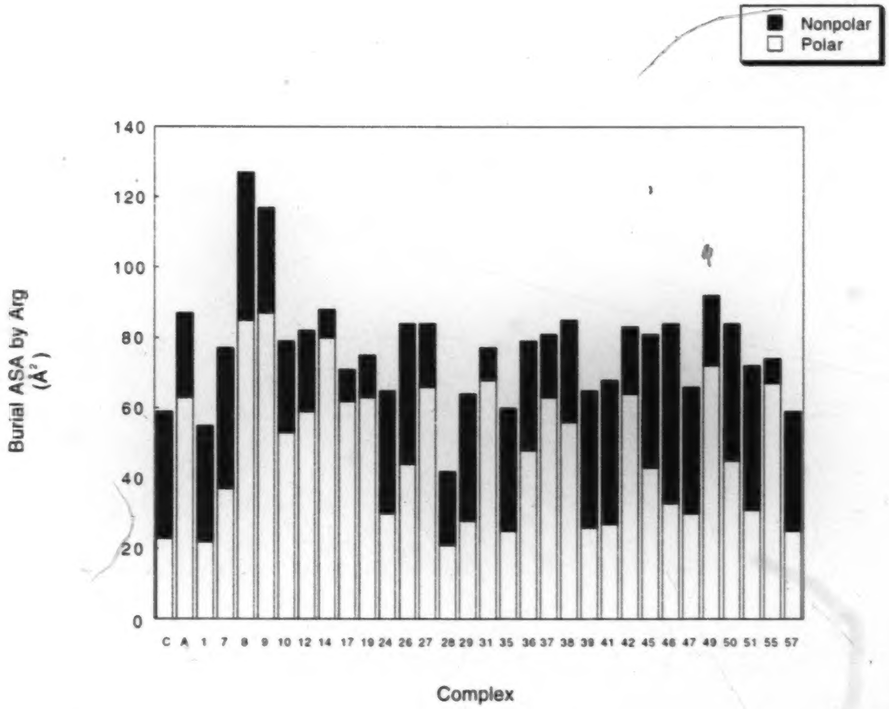


Figure 24- *Accessible surface area buried by the Arg residue of MLR in the 30 best calculated NMR complexes. The average nonpolar burial by Arg is 29 Å^2 (standard deviation = 12 Å^2, standard error = 2 Å^2), while the average polar area is 48 Å^2 (standard deviation =20 Å^2, standard error = 4 Å^2).*

Figure 25 displays the burial of the accessible surface area for the third major side chain residue in MLR which is Leu for the set of 32 high affinity complexes. The average nonpolar burial of Leu is 66 Å^2 (standard deviation = 16 Å^2, standard error = 3 Å^2), while the average polar burial is 1 Å^2 (standard deviation = 1Å^2, standard error =0 Å^2). Leu has almost exclusively nonpolar accessible surface area burial. Similar to Adda there is a wide variety in the magnitude of nonpolar burial when comparing these tight binding affinity complexes. Some of the complexes that had reduced Adda nonpolar burial have high Leu burial (e.g. complexes A average, #19, and #29) indicating that Leu may be compensating for the lack of Adda burial in those complexes. The opposite is true as well, complexes having low nonpolar Leu burial have high Adda nonpolar burial (e.g. complexes #8 and #9).

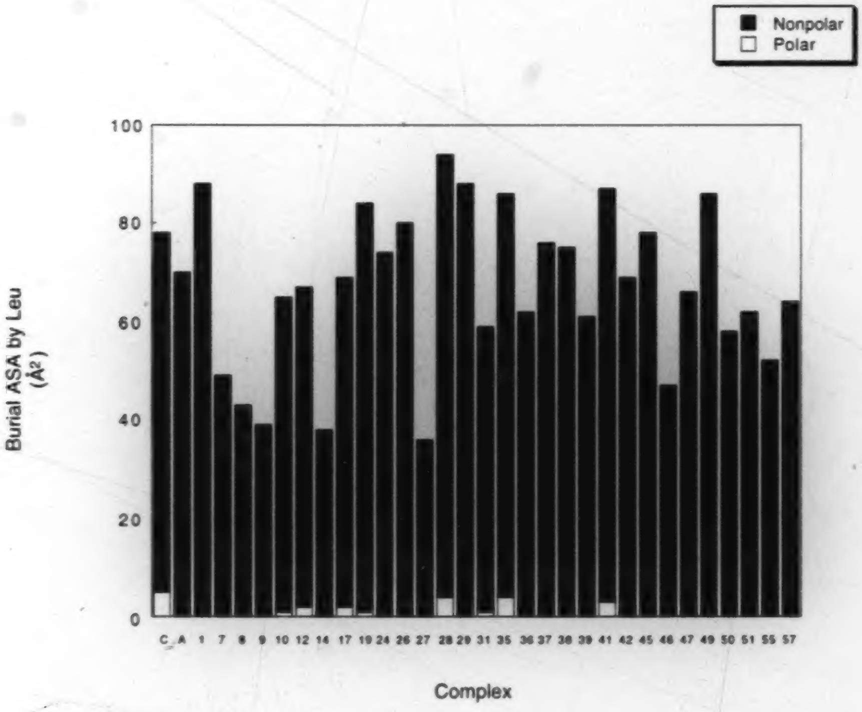


Figure 25- *Burial of the accessible surface area for the Leu residue in MLR for the set of 32 high affinity complexes. The average nonpolar burial of Leu is 66 Å^2 (standard deviation = 16 Å^2, standard error = 3 Å^2), while the average polar burial is 1 Å^2 (standard deviation = 1Å^2, standard error =0 Å^2).*

Figure 26 shows the other residues which are limited to the cyclic backbone conformation of MLR because of having only 1 rotable bond for their side chains. They are Ala, Masp, Glu (modified), and Mdha. The average nonpolar burial for Ala is 17 Å^2 (standard deviation = 8 Å^2, standard error = 2 Å^2), while the average polar burial for Ala is 3 Å^2 (standard deviation = 3 Å^2, standard error = 1 Å^2). For Masp, the average nonpolar burial is 4 Å^2 (standard deviation = 1 Å^2, standard error = 0 Å^2), while the average polar burial is 24 Å^2 (standard deviation = 4 Å^2, standard error = 1 Å^2). The average nonpolar burial for Glu is 24 Å^2 (standard deviation = 7 Å^2, standard error = 1 Å^2), while the average polar burial is higher at 63 Å^2 (standard deviation = 7 Å^2, standard error = 1 Å^2). The final backbone residue Mdha has an average nonpolar burial of 65 Å^2 (standard deviation = 14 Å^2, standard error = 2 Å^2), while the average polar burial is 21 Å^2 (standard deviation = 11 Å^2, standard error = 2 Å^2). As observed in Figure 9, the Ala and Masp residues of MLR are highly surface accessible when MLR is bound to PP-1c and therefore bury negligible area. The only major contribution is the carboxyl group of Masp which interacts with Arg 91 of PP-1c. The Glu and Mdha are on the other side of MLR which is completely interacting with PP-1c making these residues nicely buried. Mdha is the residue which covalently links with C-273 of PP-1c.

Most of the 32 tight affinity complexes have backbones residues which bury a combined total of nonpolar and polar area in excess of 200 Å^2 (Figure 26). The exceptions are complexes #10, #17, and #49. All three of these complexes have stronger than average burial of the Adda, Arg, and Leu residues (Figures 23-25) which compensates for the loss in burial that the backbones have with respect to the other complexes where the backbone has a nicer fit.

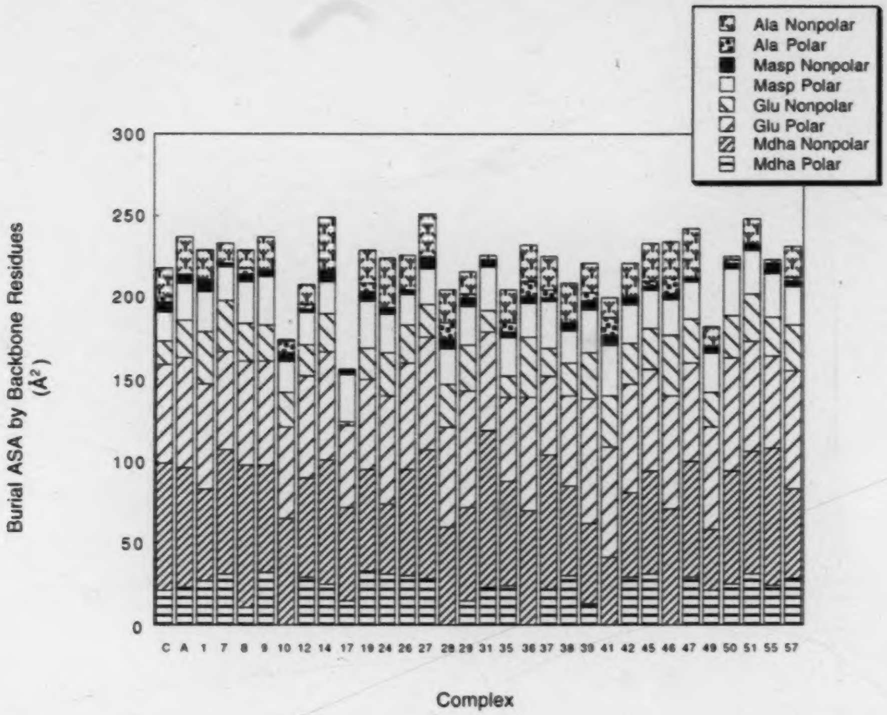


Figure 26- *Burial of accessible surface area for MLR backbone residues Ala, Masp, Glu (modified), and Mdha for the 32 high affinity complexes. The average nonpolar burial for Ala is 17 $\AA^2$ (standard deviation = 8 $\AA^2$, standard error = 2 $\AA^2$), while the average polar burial for Ala is 3 $\AA^2$ (standard deviation = 3 $\AA^2$, standard error = 1 $\AA^2$). For Masp, the average nonpolar burial is 4 $\AA^2$ (standard deviation = 1 $\AA^2$, standard error = 0 $\AA^2$), while the average polar burial is 24 $\AA^2$ (standard deviation = 4 $\AA^2$, standard error = 1 $\AA^2$). The average nonpolar burial for Glu is 24 $\AA^2$ (standard deviation = 7 $\AA^2$, standard error = 1 $\AA^2$), while the average polar burial is higher at 63 $\AA^2$ (standard deviation = 7 $\AA^2$, standard error = 1 $\AA^2$). The final backbone residue Mdha has an average nonpolar burial of 65 $\AA^2$ (standard deviation = 14 $\AA^2$, standard error = 2 $\AA^2$), while the average polar burial is 21 $\AA^2$ (standard deviation = 11 $\AA^2$, standard error = 2 $\AA^2$).*

Analysis of NMR docked MLR:PP-1c complexes:Part 3- 16 Low Affinity Complexes

Out of the 47 docked NMR complexes (including NMR average), 16 complexes had significantly less than average free energies of dissociation (Figure 12). One of the 16 complexes did not dock in the correct binding pocket as the crystal complex, and therefore was excluded from any data analysis. This section will focus on the remaining 15 complexes in terms of burial of accessible surface area of microcystin-LR residues.

First of all, the total microcystin-LR accessible surface area burial in terms of nonpolar and polar area for the bad complexes is given in Figure 27. The average nonpolar burial of accessible surface area by MLR for the low affinity calculated complexes is 365 Å^2 (standard deviation = 38 Å^2, standard error = 10 Å^2), while the average polar burial is 136 Å^2 (standard deviation = 30 Å^2, standard error = 8 Å^2). Compared to the 32 tight binding complexes in Part 2 (Figure 21), the 15 low affinity have significantly less nonpolar and polar burial of MLR based on standard error.

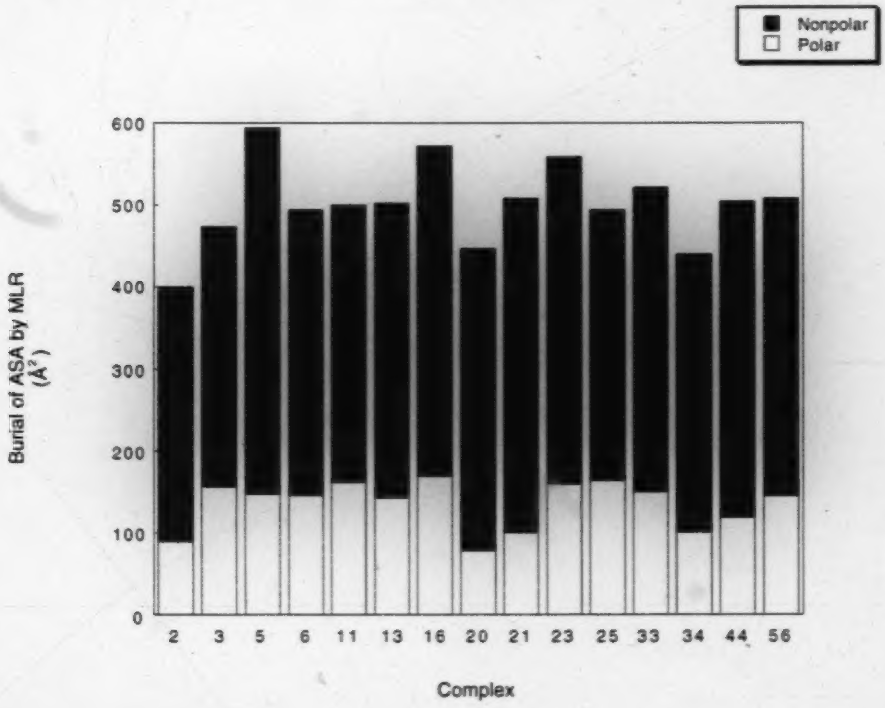


Figure 27- *Total microcystin-LR accessible surface area burial in terms of nonpolar and polar area for the 15 low affinity complexes. The average nonpolar burial of accessible surface area by MLR for the bad calculated complexes is 365 Å^2 (standard deviation = 38 Å^2, standard error = 10 Å^2), while the average polar burial is 136 Å^2 (standard deviation = 30 Å^2, standard error = 8 Å^2).*

Figure 28 is the PP-1c nonpolar and polar burial of accessible surface area for the 15 low affinity complexes. The average nonpolar burial for PP-1c is 219 Å^2 (standard deviation = 26 Å^2, standard error = 7 Å^2), while the polar burial is 244 Å^2 (standard deviation = 23 Å^2, standard error = 6 Å^2). As with MLR burial, there is significantly less nonpolar and polar area buried by PP-1c for the 15 low affinity complexes in comparison to the 32 high affinity complexes (Figure 22) based on standard error. Thus, the simple reason for why these complexes do not bind as well is the basic lack of overall nonpolar and polar surface area burial.

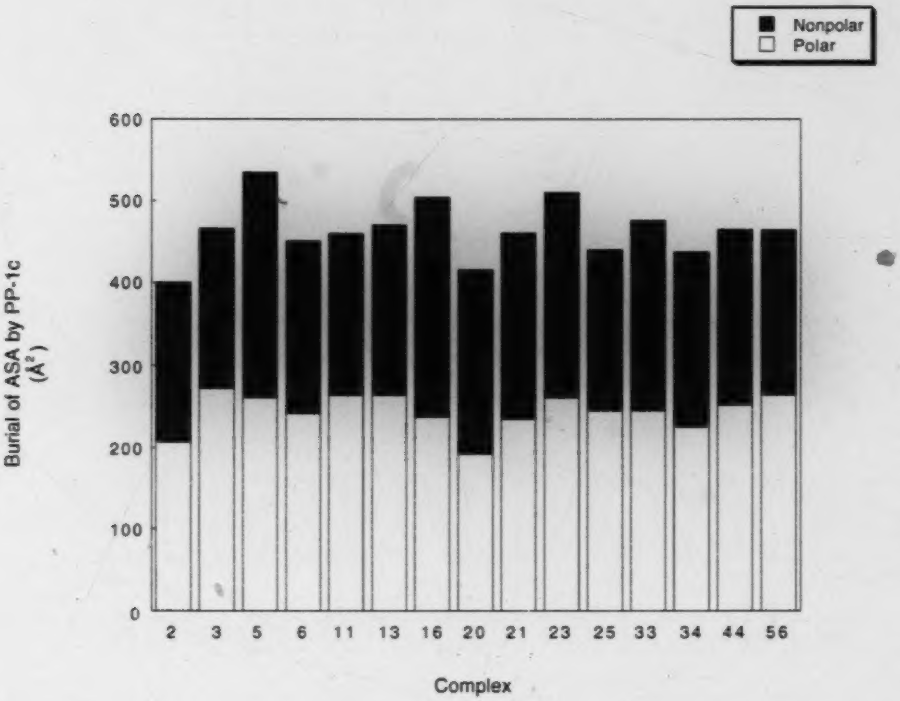


Figure 28- *PP-1c nonpolar and polar burial of accessible surface area for the 15 low affinity complexes. The average nonpolar burial for PP-1c is 219 Å^2 (standard deviation = 26 Å^2, standard error = 7 Å^2), while the polar burial is 244 Å^2 (standard deviation = 23 Å^2, standard error = 6 Å^2).*

Can we determine why the 15 low affinity complexes are burying less area? An examination of the specific residues of MLR can answer this question. Figure 29 shows the nonpolar and polar burial of accessible surface area for the Adda residue of MLR for the 15 low affinity complexes. The average nonpolar burial by Adda is 214 Å^2 (standard deviation = 52 Å^2, standard error = 14 Å^2), while the average polar burial is 20 Å^2 (standard deviation = 7 Å^2, standard error = 2 Å^2). Surprisingly, there is no significant difference in the burial of nonpolar area between the 15 low affinity complexes and the 32 high affinity complexes (Figure 23) based on standard error. There is a significant difference in the burial of polar area by Adda between the two groups of complexes, but since the total difference is only 5 Å^2 this is relatively neglible. The only significantly low burial of Adda is in complex #25, which is the probable reason for why this complex is of low affinity. However, the other low affinity complexes have adequate burial of Adda so for them there must be other reasons.

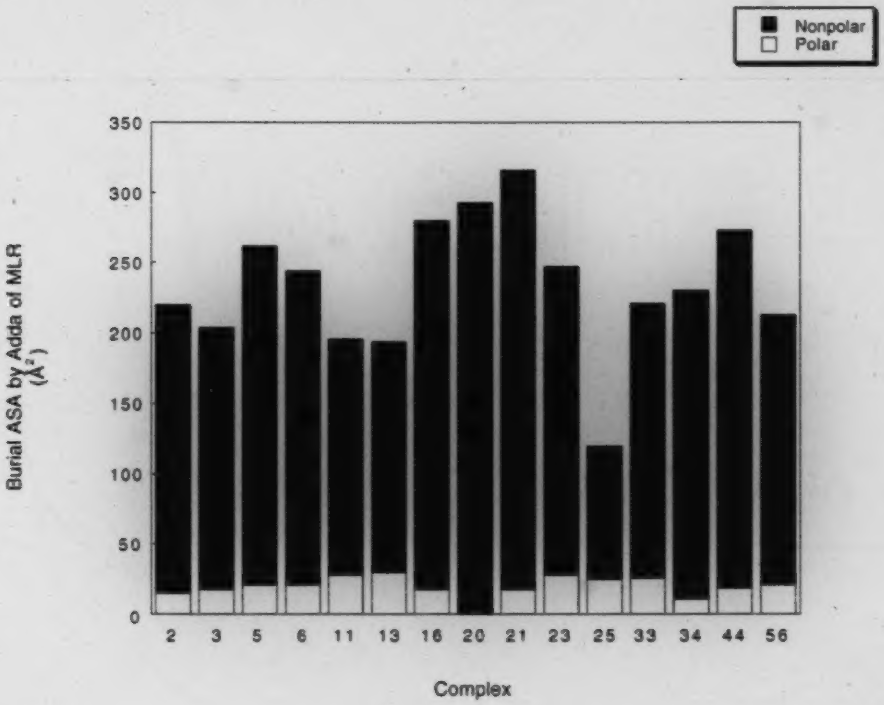


Figure 29- *Nonpolar and polar burial of accessible surface area for the Adda residue of MLR for the 15 low affinity complexes. The average nonpolar burial by Adda is 214 Å^2 (standard deviation = 52 Å^2, standard error = 14 Å^2), while the average polar burial is 20 Å^2 (standard deviation = 7 Å^2, standard error = 2 Å^2).*

Figure 30 is the burial of the accessible surface area for the Arg residue for the 15 low affinity complexes. The average nonpolar burial of Arg is 28 Å^2 (standard deviation = 14 Å^2, standard error = 4 Å^2), while the average polar burial is only 25 Å^2 (standard deviation = 10 Å^2, standard error = 3 Å^2). In comparison with the 32 high affinity complexes (Figure 24), the low affinity complexes bury significantly less polar Arg accessible surface area but bury the same nonpolar area based on standard error. Relative to the total area buried, the difference in polar Arg area can be considered to be of a minor consequence since the overall difference is only ~20 Å^2. This implies that for the low affinity complexes the Arg side chain is not generally completely properly buried. However, Figure 30 does provide an explanation of why the individual complexes #20, #34, and #44 are of low affinity binding because these complexes have significantly less burial of the Arg residue.

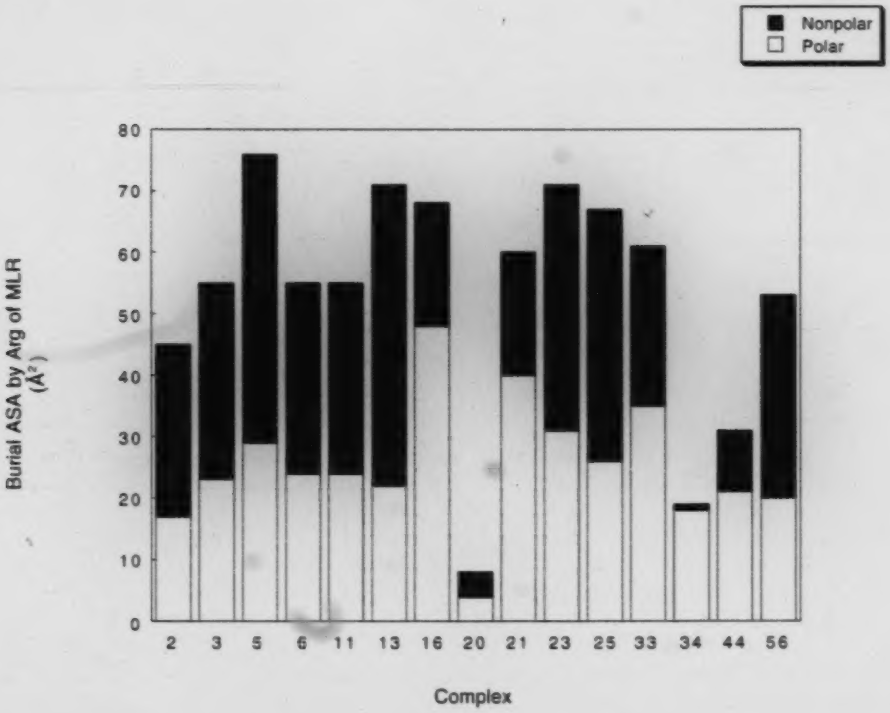


Figure 30- *Burial of the accessible surface area of the Arg residue for the 15 low affinity complexes. The average nonpolar burial for Arg is 28 Å^2 (standard deviation = 14 Å^2, standard error = 4 Å^2), while the average polar burial is only 25 Å^2 (standard deviation = 10 Å^2, standard error = 3 Å^2).*

Leu is the final major side chain residue in MLR that can have an impact on binding. Figure 31 is the burial of accessible surface area for this residue with respect to the 15 low affinity complexes. The average burial of nonpolar area for Leu is 34 Å^2 (standard deviation = 16 Å^2, standard error = 4 Å^2), while the average polar area buried is 0 Å^2 (standard deviation = 1 Å^2, standard error = 0 Å^2). Similar to Arg, Leu shows a significant difference based on standard error in comparing the 32 high affinity complexes (Figure 25) with the 15 low affinity complexes except of course in terms of nonpolar rather than polar area buried. The difference on average is of minor consequence being near a 30 Å^2 reduction in burial of nonpolar area for the low affinity complexes. There is no difference in the burial of polar area. Figure 31 shows that the reason why complexes #6, #16, and #44 have low affinity binding may be because a lack of burial of Leu.

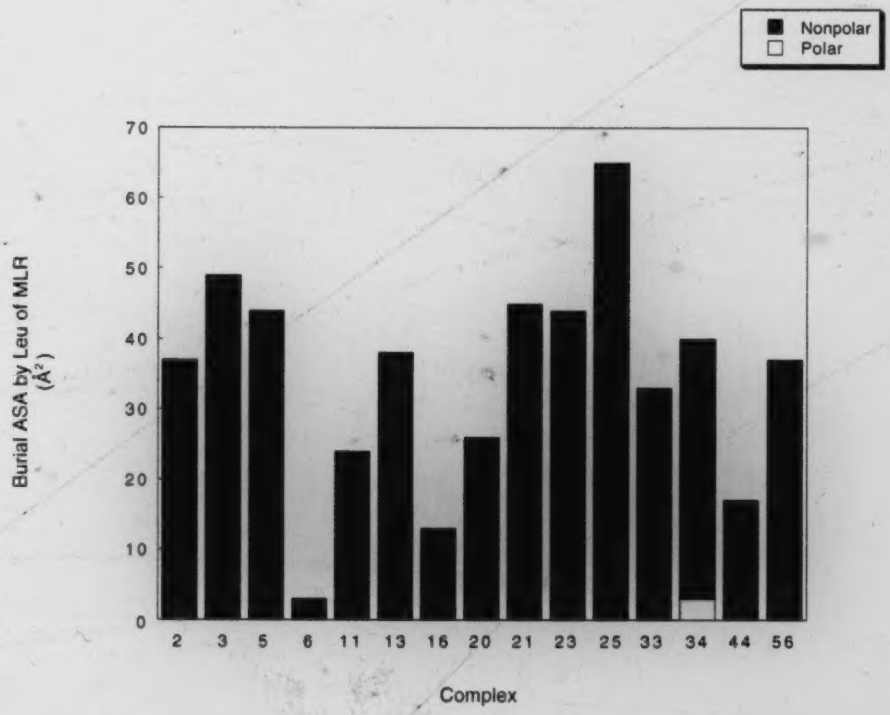


Figure 31- *Burial of accessible surface area for this residue with respect to the 15 low affinity complexes. The average burial of nonpolar area for Leu is 34 Å^2 (standard deviation = 16 Å^2, standard error = 4 Å^2), while the average polar area buried is 0 Å^2 (standard deviation = 1 Å^2, standard error = 0 Å^2).*

The last possible explanation for why the low affinity complexes have less binding capability is the burial of the backbone residues Ala, Masp, Glu, and Mdha. Figure 32 displays the burial of these backbone residues of MLR for the 15 low affinity complexes. The average nonpolar burial for Ala is 4 Å^2 (standard deviation = 6 Å^2, standard error = 4 Å^2), while the average polar burial is 2 Å^2 (standard deviation = 3 Å^2, standard error = 1 Å^2). For Masp, the average nonpolar burial is 4 Å^2 (standard deviation = 2 Å^2, standard error = 1 Å^2), while the average polar burial is 24 Å^2 (standard deviation = 7 Å^2, standard error = 2 Å^2). For Glu, the average nonpolar burial is 24 Å^2 (standard deviation = 7 Å^2, standard error = 2 Å^2), while the average polar burial is 58 Å^2 (standard deviation = 12 Å^2, standard error = 3 Å^2). The final backbone residue Mdha has an average nonpolar burial of 57 Å^2 (standard deviation = 26 Å^2, standard error = 7 Å^2), while the average polar burial is 7 Å^2 (standard deviation = 9 Å^2, standard error = 2 Å^2). Based on a standard error comparison of the averages between the 15 low affinity complexes and 32 high affinity complexes (Figure 26) a significant difference can be found in the burial of Ala nonpolar area, Glu polar area, and Mdha polar area. The magnitude of the differences are approximately 10 Å^2, 5 Å^2, and 10 Å^2 respectively. No other significant differences occurred in the other possible categories. Like the Arg polar and Leu nonpolar differences these backbone differences can be considered to be only minor. However, Figure 32 does show that complexes #2, #20, and #21 have significantly less burial of total backbone area being well below 200 Å^2, indicating that the backbone of MLR is incorrectly interacting with PP-1c.

In summary, it is clear that the 15 low affinity complexes on average have less nonpolar and polar accessible surface area buried for both PP-1c and MLR in comparison to the 32 high affinity complexes. Despite large structural/position variations in the Adda residue only complex #25 of the low affinity group had an incorrectly buried Adda residue. Complexes #6 and #16 had low Leu burial, complexes #2 and #21 had low backbone residue burial, complex #34 had low Arg burial, complex #20 had low Arg and backbone burial, and complex #44 had low Arg and Leu burial. Complexes #3, #5, #11, #13, #23, #33, and #56 had no apparent significant reduction in burial in any of the MLR residues. For these complexes the explanation of why they have low affinity binding must be a summation of minor but significant differences in their burial of the Arg, Leu, and some of the backbone residues. These minor differences must compound to give an overall reduction in surface area burial that leads to a reduction in binding affinity.

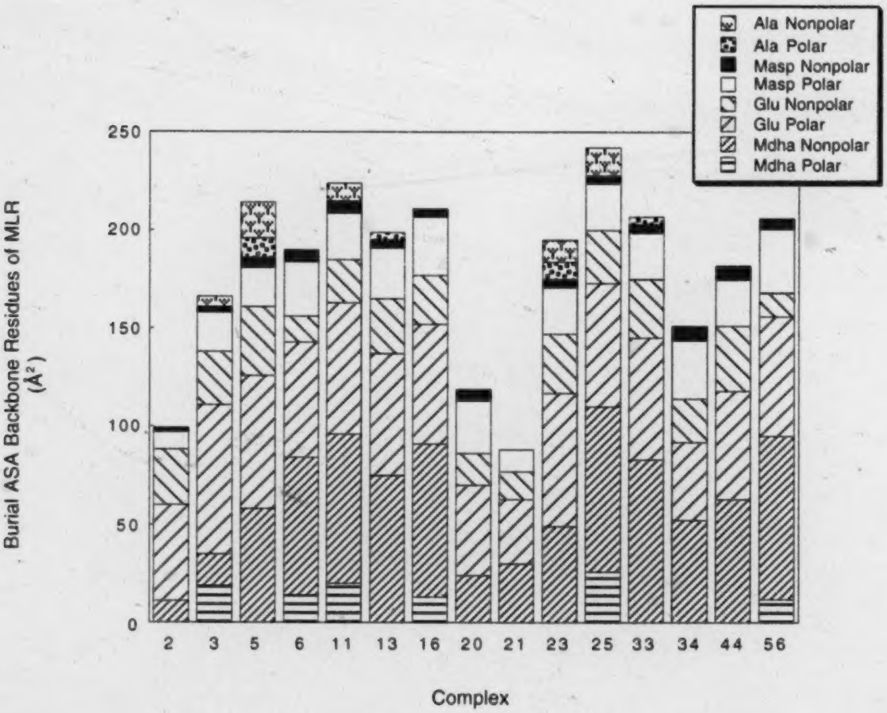


Figure 32- *Burial of backbone residues Ala, Masp, Glu, and Mdha of MLR for the 15 low affinity complexes. The average nonpolar burial for Ala is 4 $Å^2$ (standard deviation = 6 $Å^2$, standard error = 4 $Å^2$), while the average polar burial is 2 $Å^2$ (standard deviation = 3 $Å^2$, standard error = 1 $Å^2$). For Masp, the average nonpolar burial is 4 $Å^2$ (standard deviation = 2 $Å^2$, standard error = 1 $Å^2$), while the average polar burial is 24 $Å^2$ (standard deviation = 7 $Å^2$, standard error = 2 $Å^2$). For Glu, the average nonpolar burial is 24 $Å^2$ (standard deviation = 7 $Å^2$, standard error = 2 $Å^2$), while the average polar burial is 58 $Å^2$ (standard deviation = 12 $Å^2$, standard error = 3 $Å^2$). The final backbone residue Mdha has an average nonpolar burial of 57 $Å^2$ (standard deviation = 26 $Å^2$, standard error = 7 $Å^2$), while the average polar burial is 7 $Å^2$ (standard deviation = 9 $Å^2$, standard error = 2 $Å^2$).*

Analysis of NMR docked MLR:PP-1c complexes:Part 4- 7 Probable Complexes

Figures 11A/B show the probabilities of all the complexes. 7 of the complexes (6 of the open circles and the crystal structure complex at r.m.s.d. of 0) have a probability of 0.05 or higher. They are the crystal complex $P_i = 0.19$, the average complex $P_i = 0.07$, complex #9 $P_i = 0.07$, complex #19 $P_i = 0.05$, complex #36 $P_i = 0.12$, complex #38 $P_i = 0.06$, and complex #42 $P_i = 0.14$. The total of the probabilities is 70% which indicates that these 7 conformations are present the majority of the time (Figure 13 *right* structural representation). The other 30% of the time is approximately equally divided at just over 1% for each of the remaining 25 high affinity complexes (Part 2). The low affinity complexes (Part 3) have neglible probabilities. This section will take a close look at the 7 most probable complexes to determine if any patterns exist in their binding.

Figure 33 is the free energy of dissociation for the 7 most probable complexes. The average free energy of dissociation is 15.1 kcal/mol (standard deviation = 0.5 kcal/mol, standard error = 0.2 kcal/mol), which is the highest of all sets of complexes.

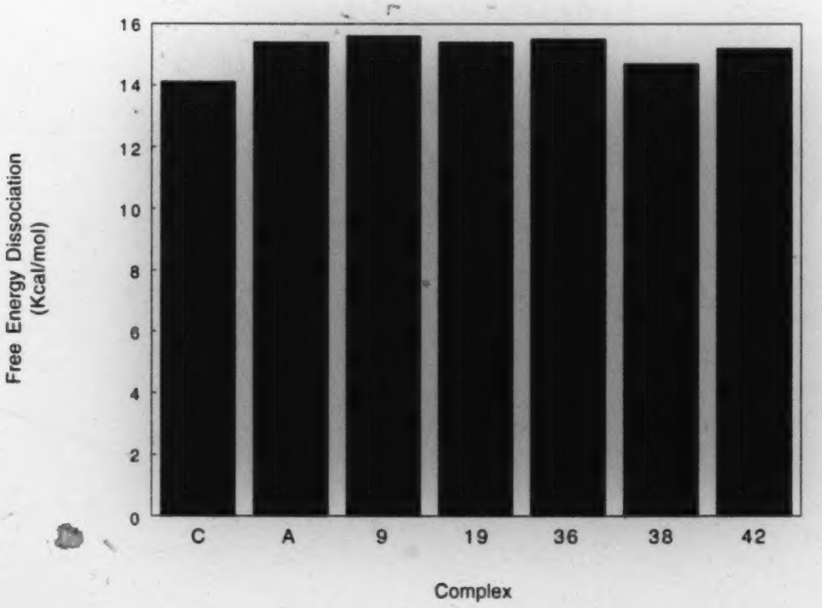


Figure 33- Free energy of dissociation for the 7 most probable complexes. The average free energy of dissociation is 15.1 kcal/mol (standard deviation = 0.5 kcal/mol, standard error = 0.2 kcal/mol).

Figure 34 is the corrected free energy of dissociation for the same 7 complexes. The average corrected free energy of dissociation is 13.7 kcal/mol (standard deviation = 0.2 kcal/mol, standard error = 0.1 kcal/mol). This corrected free energy of dissociation is correlated to P_i, with higher free energies having greater probabilities. The average is about 1 kcal/mol higher than the average of the set of 32 best complexes. The 7 essentially are the best of the best.

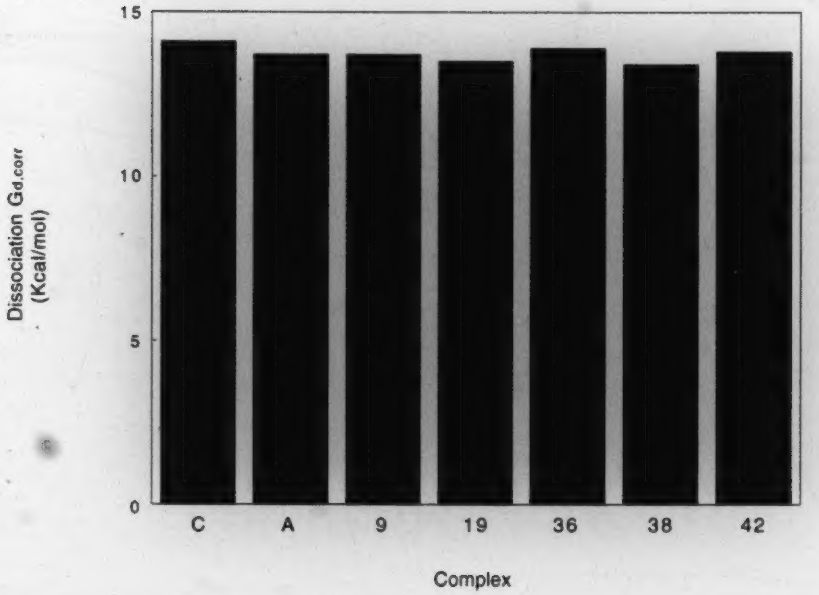


Figure 34- Corrected free energy of dissociation for the same 7 complexes. The average corrected free energy of dissociation is 13.7 kcal/mol (standard deviation = 0.2 kcal/mol, standard error = 0.1 kcal/mol).

Figure 35 is the dissociation entropy and enthalpy for the 7 highest probability complexes. The average entropy value is 11.3 kcal/K/mol (standard deviation = 3.1 kcal/K/mol, standard error = 1.2 kcal/K/mol), while the average enthalpy is much lower at 3.8 kcal/mol (standard deviation = 3.4 kcal/mol, standard error = 1.3 kcal/mol). This is another fine example of the entropy-enthalpy compensation effect, where complexes of similar free energies have different entropy and enthalpy contributions dependent upon small changes in structure/position of the ligand MLR with respect to PP-1c which result in variable changes in accessible surface area burial. As with the other high affinity complexes (Part 2), the greatest probability complexes are mostly dependent upon entropy for their binding. However, complexes A (average), #19, and #42 have strong enthalpy contributions.

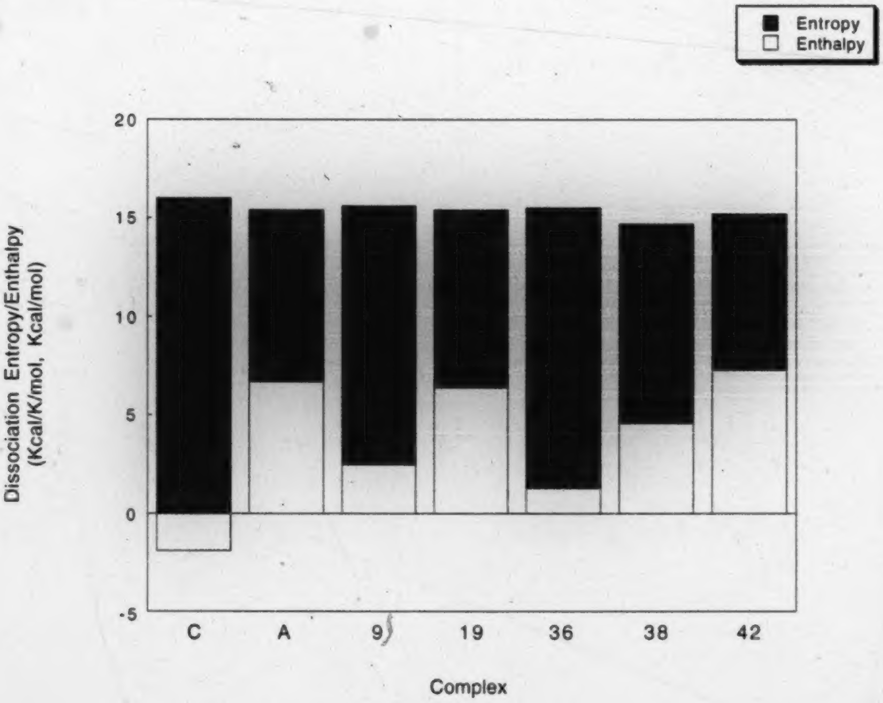


Figure 35- *Dissociation entropy and enthalpy for the 7 highest probability complexes. The average entropy value is 11.3 kcal/K/mol (standard deviation = 3.1 kcal/K/mol, standard error = 1.2 kcal/K/mol), while the average enthalpy is much lower at 3.8 kcal/mol (standard deviation = 3.4 kcal/mol, standard error = 1.3 kcal/mol).*

Figure 36 displays the nonpolar and polar burial of accessible surface area of MLR when inhibiting PP-1c for the 7 most probable complexes. The average nonpolar burial of MLR is 436 Å^2 (standard deviation = 78 Å^2, standard error = 29 Å^2), while the average polar burial is 194 Å^2 (standard deviation = 29 Å^2, standard error = 11 Å^2). There is no consistent pattern within this set of complexes in terms of amount buried. The only conclusion is that on average twice as much nonpolar area is buried than polar area by microcystin-LR, but there is a large variety within that conclusion. The burial in this group of complexes is not significantly different from the 32 high affinity set of complexes (Figure 21). It is interesting to note that the largest nonpolar burial is in complexes C (crystal), complex #9, and complex #36.

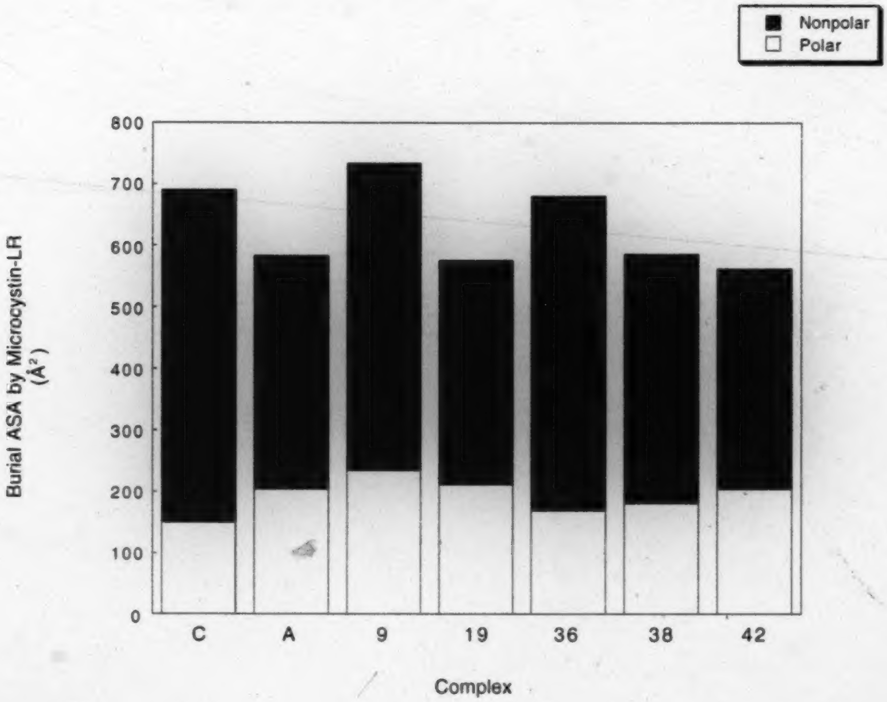


Figure 36- *Nonpolar and polar burial of accessible surface area of MLR when inhibiting PP-1c for the 7 most probable complexes. The average nonpolar burial of MLR is 436 Å^2 (standard deviation = 78 Å^2, standard error = 29 Å^2), while the average polar burial is 194 Å^2 (standard deviation = 29 Å^2, standard error = 11 Å^2).*

The corresponding burial of accessible surface area by PP-1c within the most probable 7 complexes is shown in Figure 37. The average nonpolar burial of accessible surface area by PP-1c is 253 Å^2 (standard deviation = 27 Å^2, standard error = 10 Å^2), while the polar burial is higher at 313 Å^2 (standard deviation =15 Å^2, standard error = 6 Å^2). There is no major significant difference in the average burial of this set of complexes in comparison to the 32 high affinity complexes. The variability in burial is somewhat reduced in comparison to MLR.

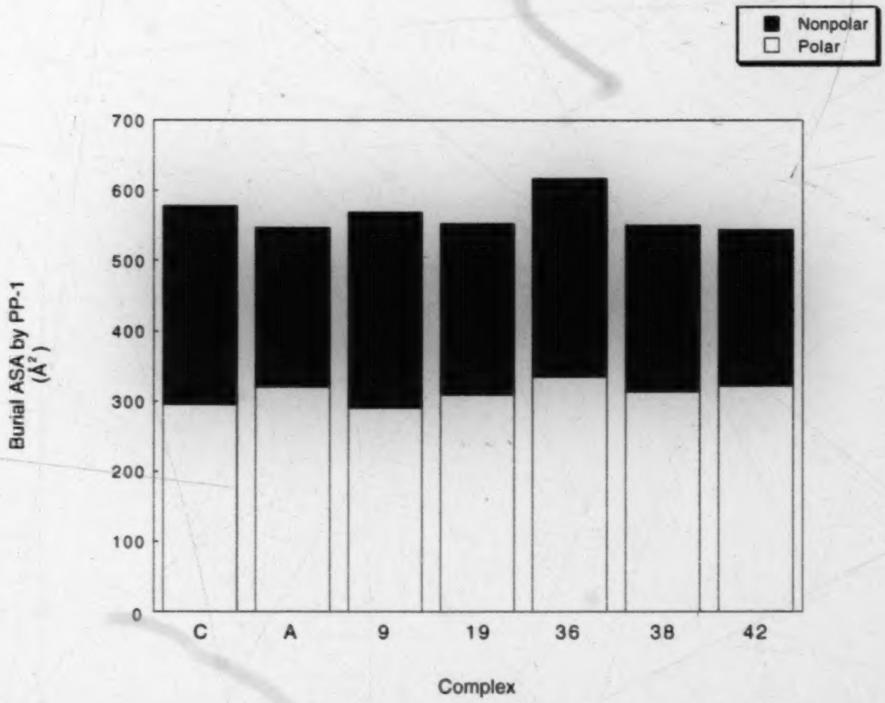


Figure 37- *Burial of accessible surface area by PP-1c within the most probable 7 complexes. The average nonpolar burial of accessible surface area by PP-1c is 253 Å^2 (standard deviation = 27 Å^2, standard error = 10 Å^2), while the polar burial is higher at 313 Å^2 (standard deviation =15 Å^2, standard error = 6 Å^2).*

Figure 38 evaluates the burial of the Adda residue of MLR in the 7 most probable complexes. The average nonpolar burial of the Adda residue is 229 Å^2 (standard deviation = 74 Å^2, standard error = 28 Å^2), while the average polar burial is 23 Å^2 (standard deviation = 3 Å^2, standard error = 1 Å^2). As observed in Figure 9 (*right*) the 7 most probable complexes ended up have two sets of conformations to the Adda side chain. The first set includes the crystal complex, complex #9, and complex #36, the second set includes complexes #19, #38, and #42, and the average complex has an Adda side chain orientation in between the two main orientations. Figure 38 supports the fact that there are two main orientations of the Adda side chain having the highest probabilities. It is clear that complexes C, #9, and #36 bury similar amounts of area, while the same is true for complexes #19, #38, and #42 that have a reduced amount of nonpolar area buried. Taking into account the actual probabilities and adding them for each set, the crystal set has P_i = 38% verses #19/#38/#42 has P_i =25%. This would agree with the fact that the dominant Adda orientation is the observed crystal structure orientation[9].

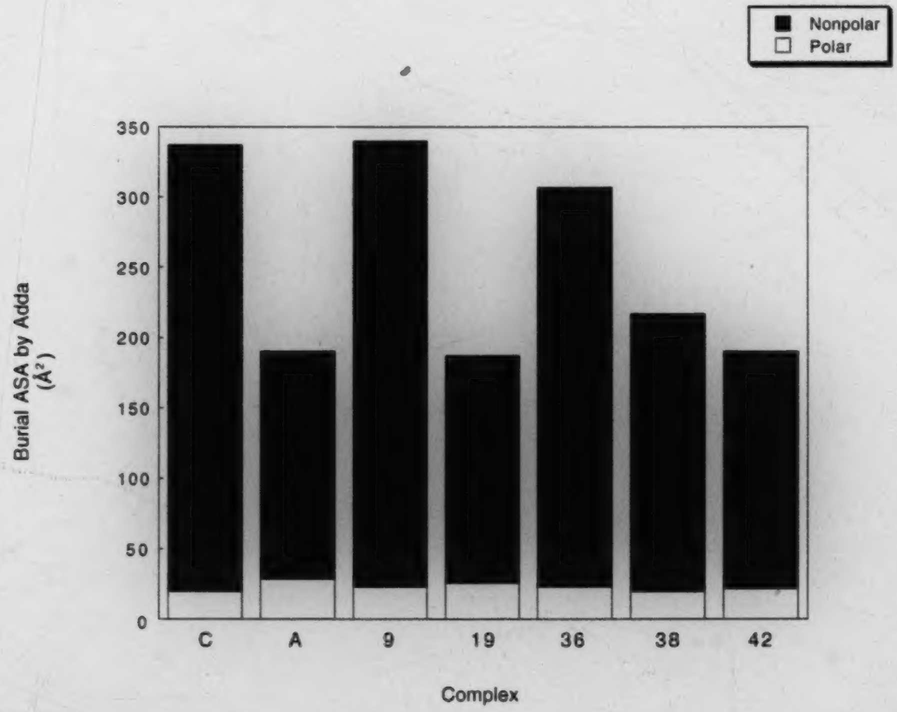


Figure 38- *Burial of the Adda residue of MLR in the 7 most probable complexes. The average nonpolar burial of the Adda residue is 229 Å^2 (standard deviation = 74 Å^2, standard error = 28 Å^2), while the average polar burial is 23 Å^2 (standard deviation = 3 Å^2, standard error = 1 Å^2).*

Unlike the Adda side chain orientation which can be distinguished into two main conformations in the 7 most probable complexes, the Arg side chain orientation cannot be grouped into discernable conformations (Figure 9 *right*). Figure 39 bears this out by displaying the nonpolar and polar burial of the Arg residue for the 7 most probable complexes. The average nonpolar burial for Arg is 26 Å^2 (standard deviation = 8 Å^2, standard error = 3 Å^2), while the average polar burial is larger at 58 Å^2 (standard deviation = 19 Å^2, standard error = 7 Å^2). There is a high degree of variability in Arg side chain burial with no discernable pattern existing. This is in agreement with Goldberg et al. (1995) where this side chain was unobservable having no main orientation. Thus, when MLR binds to PP-1c, Arg has no loss in conformational flexibility while Adda has a greatly reduced conformational flexibility. This residue can be switched to any other standard amino acid in the microcystin family of toxins.

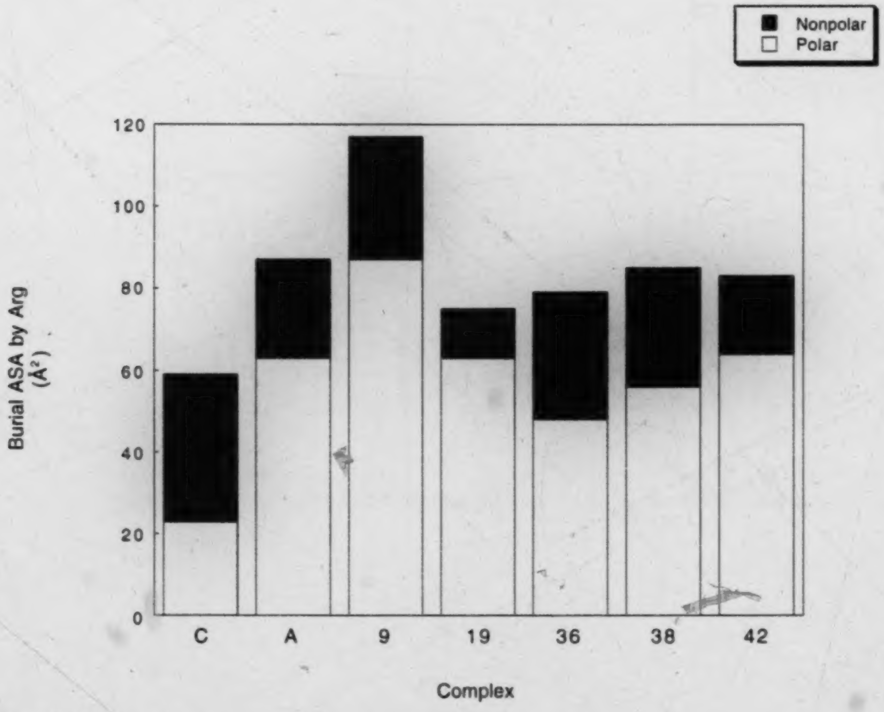


Figure 39- *Nonpolar and polar burial of the Arg residue for the 7 most probable complexes. The average nonpolar burial for Arg is 26 Å^2 (standard deviation = 8 Å^2, standard error = 3 Å^2), while the average polar burial is larger at 58 Å^2 (standard deviation = 19 Å^2, standard error = 7 Å^2).*

Figure 40 displays the Leu residue of MLR in terms of nonpolar and polar burial of accessible surface area for the 7 most probable conformations. The average nonpolar burial of Leu is 67 Å^2 (standard deviation = 14 Å^2, standard error = 5 Å^2), while the average polar burial is 1 Å^2 (standard deviation = 2 Å^2, standard error = 1 Å^2). No discernible pattern exists for Leu burial within this group of complexes. In fact, complex #9 has a rather low burial of this residue. This residue can be switched to any other standard amino acid and is not thought to be critical for activity.

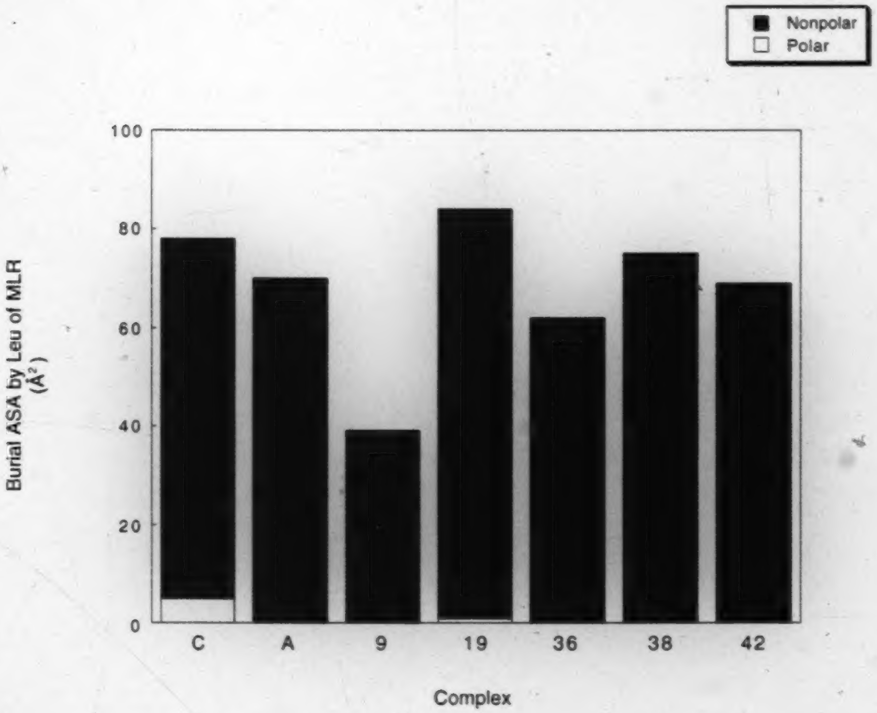


Figure 40- *Nonpolar and polar burial of accessible surface area for the Leu residue of MLR in the 7 most probable conformations. The average nonpolar burial of Leu is 67 Å^2 (standard deviation = 14 Å^2, standard error = 5 Å^2), while the average polar burial is 1 Å^2 (standard deviation = 2 Å^2, standard error = 1 Å^2).*

Figure 41 has the burial of nonpolar and polar burial of accessible surface area for the backbone residues Ala, Masp, Glu, and Mdha for the 7 most probable complexes. The average nonpolar burial for Ala is 21 Å^2 (standard deviation = 2 Å^2, standard error = 1 Å^2), while the polar burial is only 2 Å^2 (standard deviation = 4 Å^2, standard error = 1 Å^2). For Masp, the average nonpolar burial is 5 Å^2 (standard deviation = 1 Å^2, standard error = 0 Å^2), while the average polar burial is 24 Å^2 (standard deviation = 5 Å^2, standard error = 2 Å^2). For Glu, the average nonpolar burial is 23 Å^2 (standard deviation = 7 Å^2, standard error = 3 Å^2), while the average polar burial is 62 Å^2 (standard deviation = 6 Å^2, standard error = 2 Å^2). The covalently linked Mdha has an average nonpolar burial of 65 Å^2 (standard deviation = 9 Å^2, standard error = 4 Å^2), while the average polar burial is 24 Å^2 (standard deviation = 11 Å^2, standard error = 4 Å^2). Figure 41 shows the necessity of having a stong burial of accessible surface area by the cyclic backbone of MLR. All 7 most probable complexes have over 200 Å^2 total backbone residue burial. The pattern of burial is highly similar when comparing members of this group with the exception of complex #36 which lacks Mdha polar burial. Therefore, a rigid less conformational flexibility backbone is a key feature to MLR inhibiting PP-1c.

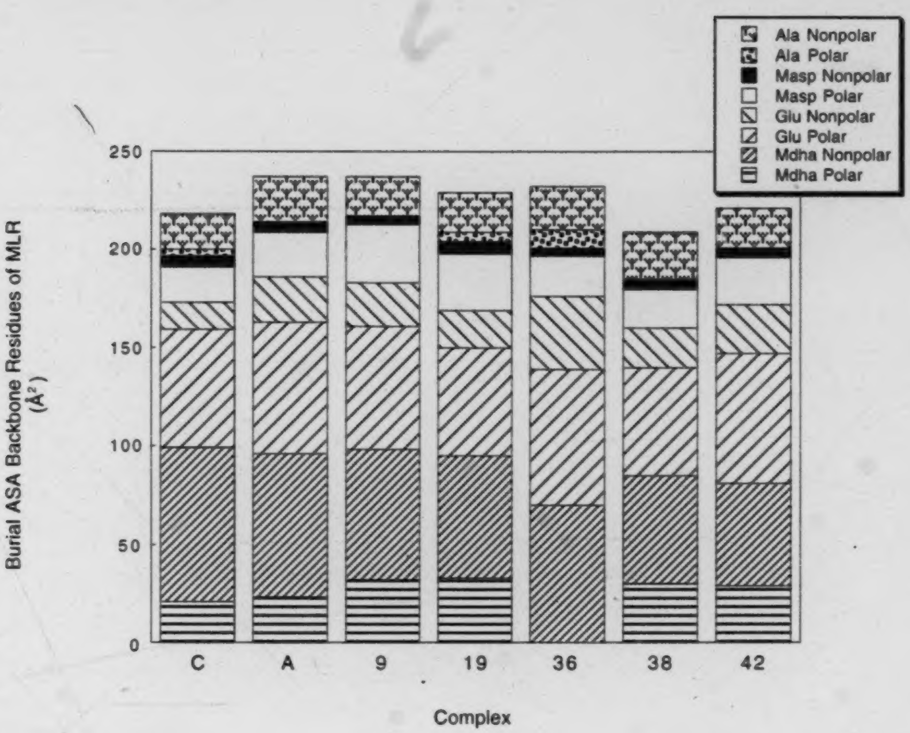


Figure 41- *Burial of nonpolar and polar burial of accessible surface area for the backbone residues Ala, Masp, Glu, and Mdha for the 7 most probable complexes. The average nonpolar burial for Ala is 21 $Å^2$ (standard deviation = 2 $Å^2$, standard error = 1 $Å^2$), while the polar burial is only 2 $Å^2$ (standard deviation = 4 $Å^2$, standard error = 1 $Å^2$). For Masp, the average nonpolar burial is 5 $Å^2$ (standard deviation = 1 $Å^2$, standard error = 0 $Å^2$), while the average polar burial is 24 $Å^2$ (standard deviation = 5 $Å^2$, standard error = 2 $Å^2$). For Glu, the average nonpolar burial is 23 $Å^2$ (standard deviation = 7 $Å^2$, standard error = 3 $Å^2$), while the average polar burial is 62 $Å^2$ (standard deviation = 6 $Å^2$, standard error = 2 $Å^2$). The covalently linked Mdha has an average nonpolar burial of 65 $Å^2$ (standard deviation = 9 $Å^2$, standard error = 4 $Å^2$), while the average polar burial is 24 $Å^2$ (standard deviation = 11 $Å^2$, standard error = 4 $Å^2$).*

Analysis of NMR docked MLL:PP-1c complexes: Part 1-All complexes

As with MLR, 56 calculated free solution NMR structures of MLL along with the average solution structure[8] were docked onto the crystal structure of PP-1c[9]. The resulting complexes were then analyzed by STC. Figure 42 displays the calculated free energy of dissociation for every MLL:PP-1c complex. The average free energy of dissociation for every MLL:PP-1c complex was 12.2 kcal/mol (standard deviation = 1.6 kcal/mol, standard error = 0.2 kcal/mol). In comparison to the MLR:PP-1c complexes (Figure 12) which had an average free energy of dissociation of 13.3 kcal/mol (standard error = 0.3 kcal/mol), MLL:PP-1c complexes had on average a significantly lower free energy of dissociation based on standard error by 0.6 kcal/mol. This calculated difference is of a rather low magnitude and can be viewed as agreeing with the experimental evidence that MLR and MLL have roughly the same binding affinity.

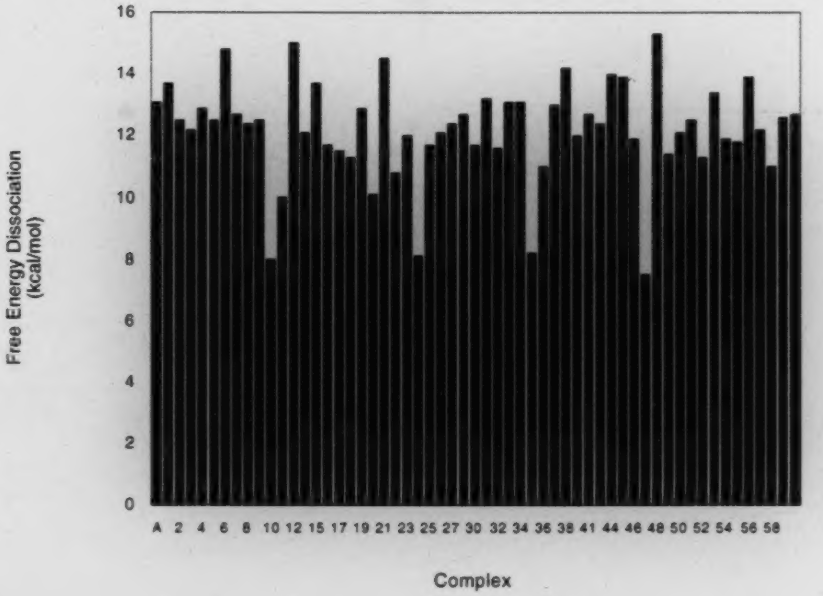


Figure 42- *Calculated free energy of dissociation for every MLL:PP-1c complex. The average free energy of dissociation for every MLL:PP-1c complex was 12.2 kcal/mol (standard deviation = 1.6 kcal/mol, standard error = 0.2 kcal/mol).*

Analysis of NMR docked MLL:PP-1c complexes:Part 2-37 High Affinity Complexes

Out of the total 57 docked MLL:PP-1c complexes, 37 had a free energy of dissociation of at least 12 kcal/mol indicating that these complexes had average or above average free energy of dissociation values based on standard error. In part 2 of this section, these high affinity complexes are studied in terms of thermodynamic calculations and burial of surface areas in the same manner as the high affinity complexes of MLR:PP-1c. The structures of the docked conformations for the 37 high affinity MLL:PP-1c complexes are shown in Figure 43.

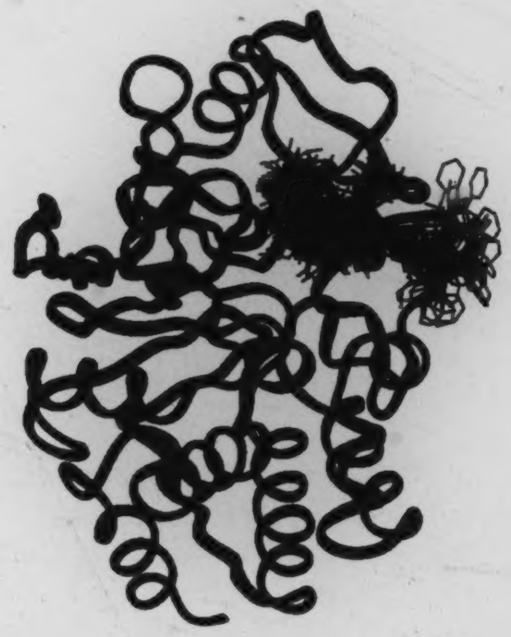


Figure 43- *Docked conformations for the 37 high affinity MLL:PP-1c complexes. Backbone red ribbon is PP-1c, while the 37 high affinity solution structures docked to PP-1c are in green.*

Figure 44 displays the free energy of dissociation for the 37 high affinity MLL:PP-1c complexes. The average free energy of dissociation is 13.1 kcal/mol (standard deviation = 0.9 kcal/mol, standard error = 0.1 kcal/mol). In comparison to the high affinity MLR:PP-1c complexes which had an average free energy of dissociation of 14.3 kcal/mol (standard error = 0.1 kcal/mol), the high affinity MLL:PP-1c complexes have a significantly less average free energy of dissociation based on standard error of 1 kcal/mol. Thus, the substitution of Leu for an Arg did have an overall impact on burial of accessible surface area contributing to a slightly lower calculated binding affinity for MLL in comparison to MLR.

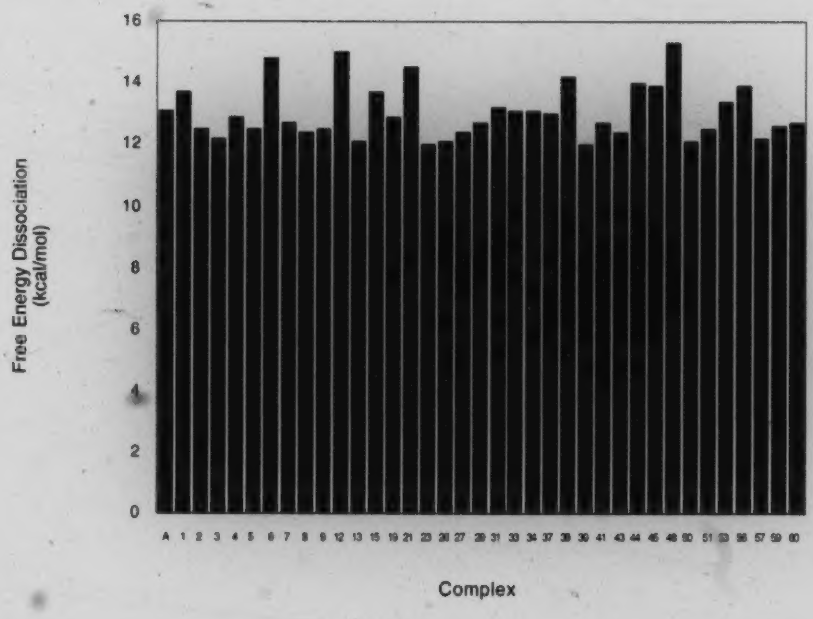


Figure 44- *Free energy of dissociation for the 37 high affinity MLL:PP-1c complexes. The average free energy of dissociation is 13.1 kcal/mol (standard deviation = 0.9 kcal/mol, standard error = 0.1 kcal/mol). The A complex represents the average solution structure of MLL docked to PP-1c for this and every other MLL figure.*

Figure 45 is the dissociation entropy and enthalpy contributions for the 37 high affinity MLL:PP-1c complexes. The average dissociation entropy is 12.1 kcal/K/mol (standard deviation = 2.3 kcal/K/mol, standard error = 0.4 kcal/K/mol), while the average dissociation enthalpy is 0.9 kcal/mol (standard deviation = 2.3 kcal/mol, standard error = 0.4 kcal/mol). As with MLR:PP-1c, the MLL:PP-1c complexes illustrate the entropy-enthalpy compensation effect where similar free energy states have varying entropy and enthalpy contributions due to minor structural and positional differences in MLL when bound to PP-1c causing different accessible surface area burial patterns. When the contribution by entropy is greater there tends to be less of a contribution by enthalpy and vice-versa. While there is no signficant difference in the entropic contribution for MLR:PP-1c (Figure 20, average dissociation entropy is 11.4 kcal/K/mol, standard error = 0.3 kcal/K/mol) verses MLL:PP-1c high affinity complexes, there is substantially less of a enthalpic contribution for the MLL:PP-1c complexes in comparison to the MLR:PP-1c (Figure 20, average dissociation enthalpy is 2.9 kcal/mol, standard error = 0.4 kcal/mol). The change from a hydrophilic Arg to a hydrophobic Leu had a dramatic reduction in favorable binding entropy and actually accounts for the overall reduction in dissociation free energy for the MLL:PP-1c complexes.

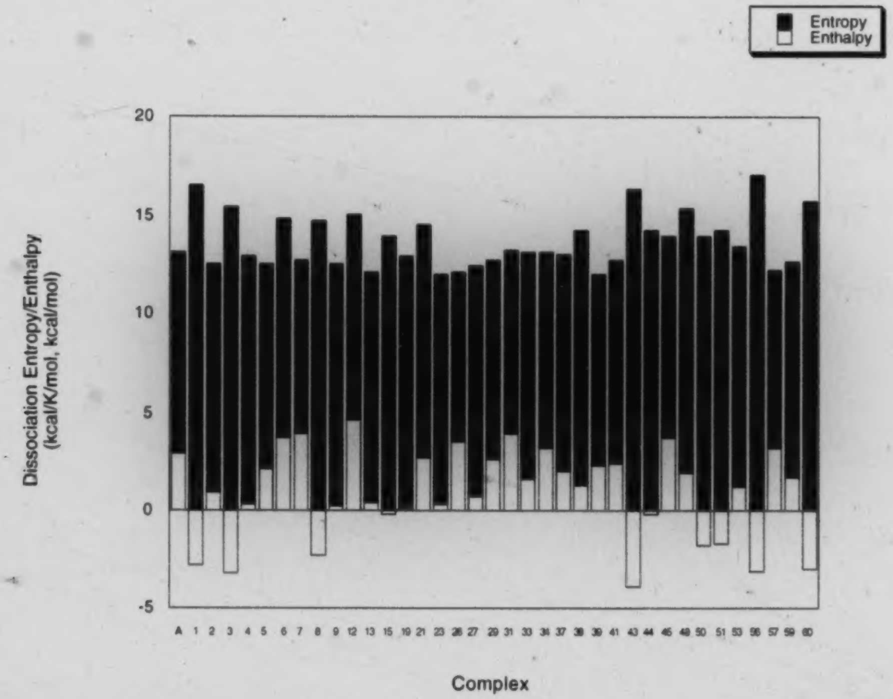


Figure 45- *Dissociation entropy and enthalpy contributions for the 37 high affinity MLL:PP-1c complexes. The average dissociation entropy for this set of complexes is 12.1 kcal/K/mol (standard deviation = 2.3 kcal/K/mol, standard error = 0.4 kcal/K/mol), while the average dissociation enthalpy is 0.9 kcal/mol (standard deviation = 2.3 kcal/mol, standard error = 0.4 kcal/mol).*

Figure 46 is the burial of nonpolar and polar accessible surface area by MLL for the 37 high affinity MLL:PP-1c complexes. The average nonpolar burial of accessible surface area is 420 Å^2 (standard deviation = 61 Å^2, standard error = 10 Å^2), while the average polar burial is 154 Å^2 (standard deviation = 22 Å^2, standard error = 4 Å^2). The high affinity MLL:PP-1c complexes had almost identical nonpolar burial in comparison to their MLR:PP-1c counterparts (Figure 21, average burial of nonpolar area is 420 Å^2, standard error = 9 Å^2). However, there was significantly less burial of polar area in MLL:PP-1c high affinity complexes in comparison to high affinity MLR:PP-1c complexes (Figure 21, average burial of polar area is 185 Å^2, standard error is 4 Å^2) based on standard error. The reduction in burial of polar surface area in MLL:PP-1c complexes accounts for less binding enthalpic contributions and a reduction in binding affinity in comparison to MLR:PP-1c high affinity complexes.

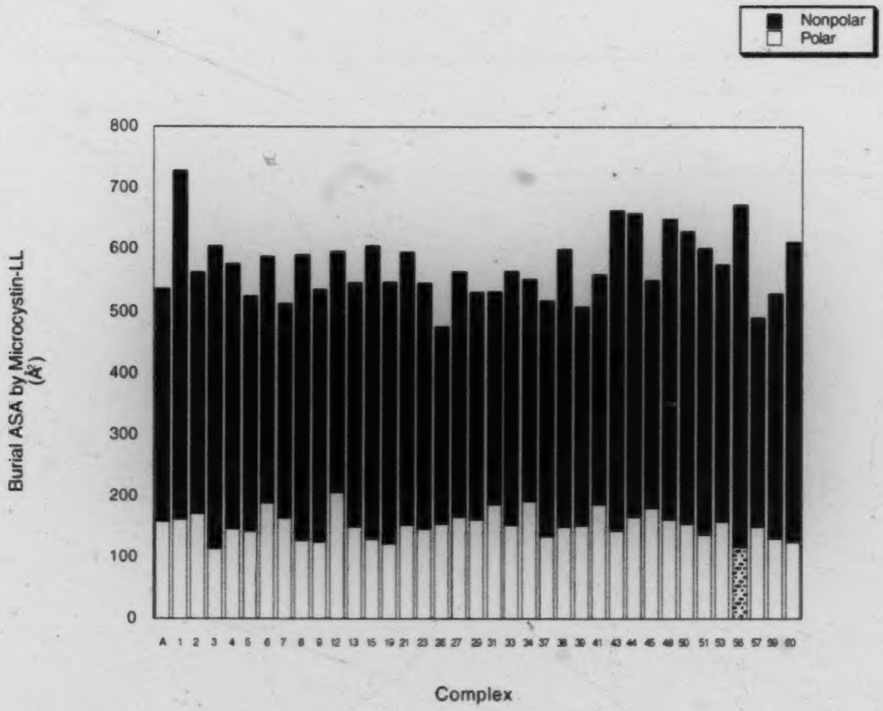


Figure 46- *Burial of nonpolar and polar ASA by MLL for the 37 high affinity MLL:PP-1c complexes. The average nonpolar burial of ASA is 420 Å^2 (standard deviation = 61 Å^2, standard error = 10 Å^2), while the average polar burial is 154 Å^2 (standard deviation = 22 Å^2, standard error = 4 Å^2).*

Figure 47 is the corresponding burial of nonpolar and polar accessible surface area by PP-1c in the 37 high affinity MLL:PP-1c complexes. The average burial of nonpolar accessible surface area by PP-1c is 238 Å^2 (standard deviation = 24 Å^2, standard error = 4 Å^2), while the average polar burial is 263 Å^2 (standard deviation = 25 Å^2, standard error = 4 Å^2). Both nonpolar and polar area buried by PP-1c are significantly reduced in MLL:PP-1c complexes relative to MLR:PP-1c high affinity complexes (Figure 22, average burial of nonpolar area by PP-1c is 257 Å^2, standard error = 5 Å^2 while the average burial of polar area is 292 Å^2, standard error = 4 Å^2) based on standard error. The magnitude of the differences are rather small and do not contribute to any signficant differences.

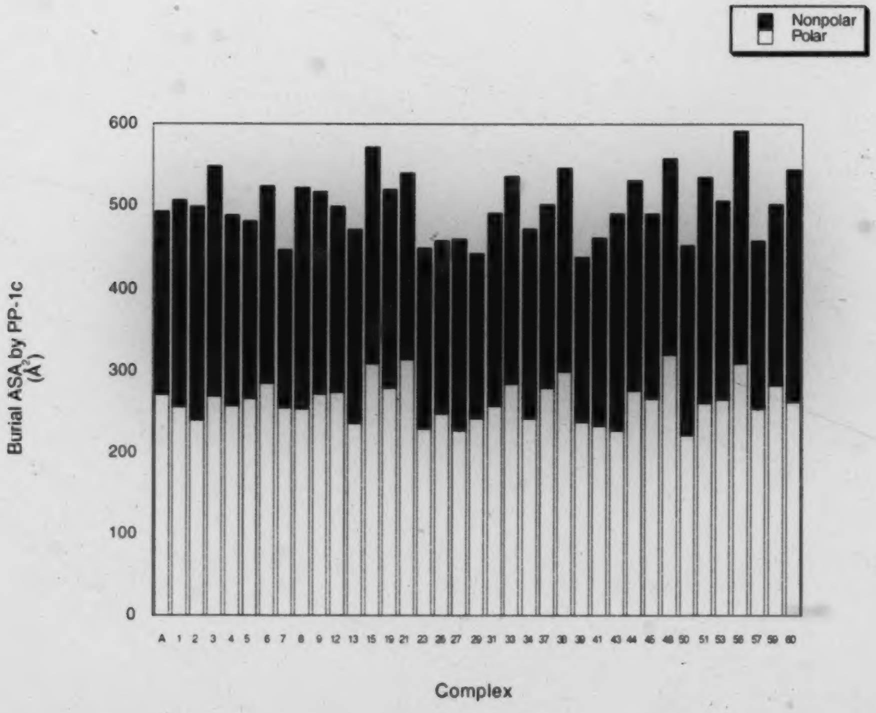


Figure 47- *Burial of nonpolar and polar accessible surface area by PP-1c in the 37 high affinity MLL:PP-1c complexes. The average burial of nonpolar accessible surface area by PP-1c is 238 Å^2 (standard deviation = 24 Å^2, standard error = 4 Å^2), while the average polar burial is 263 Å^2 (standard deviation = 25 Å^2, standard error = 4 Å^2).*

Figure 48 is the burial of the Adda residue of MLL for the 37 high affinity MLL:PP-1c complexes. The average burial of the nonpolar accessible surface area by Adda is 219 Å^2 (standard deviation = 53 Å^2, standard error = 9 Å^2), while the average polar burial is 26 Å^2 (standard deviation = 3 Å^2, standard error = 1 Å^2). Compared to high affinity MLR:PP-1c complexes (Figure 23, average nonpolar burial of accessible surface area by Adda is 216 Å^2, standard error = 25Å^2, while the polar burial by Adda is only 25 Å^2, standard error = 1 Å^2), the Adda residue for MLL:PP-1c complexes have almost identical averages for both nonpolar and polar burial of acessible surface area. Not only are the averages similar but both sets of complexes have a large variability in burial of nonpolar area as illustrated in the figure below.

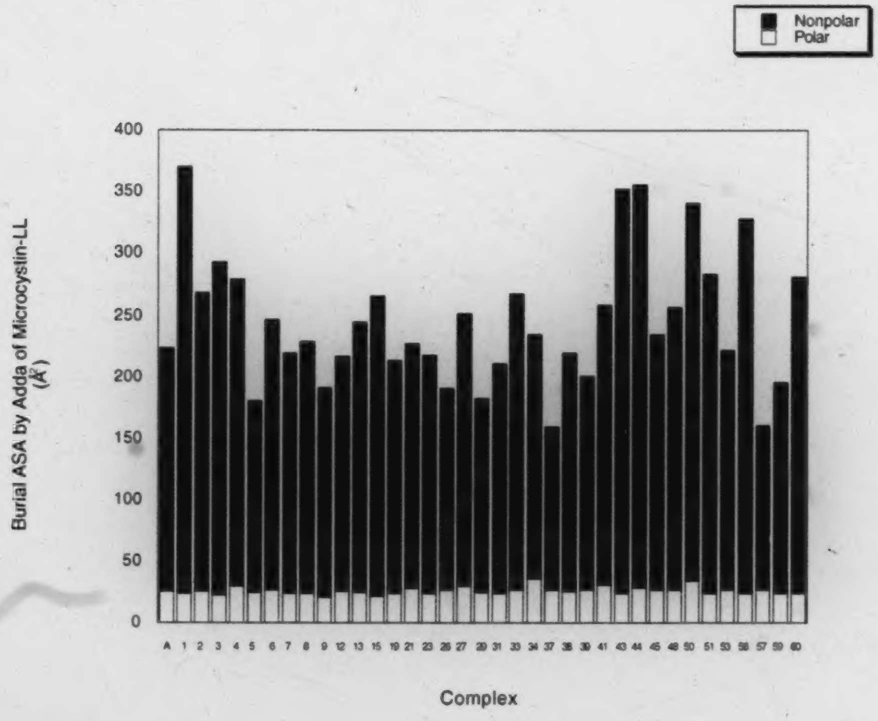


Figure 48- *Burial of the Adda residue of MLL for the 37 high affinity MLL:PP-1c complexes. The average burial of the nonpolar accessible surface area by Adda is 219 Å^2 (standard deviation = 53 Å^2, standard error = 9 Å^2), while the average polar burial is 26 Å^2 (standard deviation = 3 Å^2, standard error = 1 Å^2).*

Figure 49 illustrates the burial of the Leu residue of MLL that replaced the Arg of MLR for the high affinity MLL:PP-1c complexes. The average burial of nonpolar accessible surface area by the new Leu is 32 Å^2 (standard deviation = 12 Å^2, standard error = 2 Å^2), while burial of polar area is 21 Å^2 (standard deviation = 3 Å^2, standard error= 1 Å^2). Replacement of Leu for Arg in MLL verses MLR is the only structural difference between the two toxins. When comparing burial of these residues between the two toxins (Arg in MLR:PP-1c is Figure 24, average nonpolar burial by Arg is 29 Å^2, standard error = 2 Å^2, while the average polar area is 48 Å^2, standard error = 4 Å^2) there is no significant difference in the burial of nonpolar accessible surface area based on standard error. However, on average the MLL:PP-1c bury significantly smaller amounts of polar accessible surface area of approximately 20 Å^2 than their MLR:PP-1c high affinity counterparts. This is to be expected due to switching a highly polar residue in MLR for a highly nonpolar residue in MLL and the Leu. There is also a high degree of variability in the nonpolar burial for the new Leu residue.

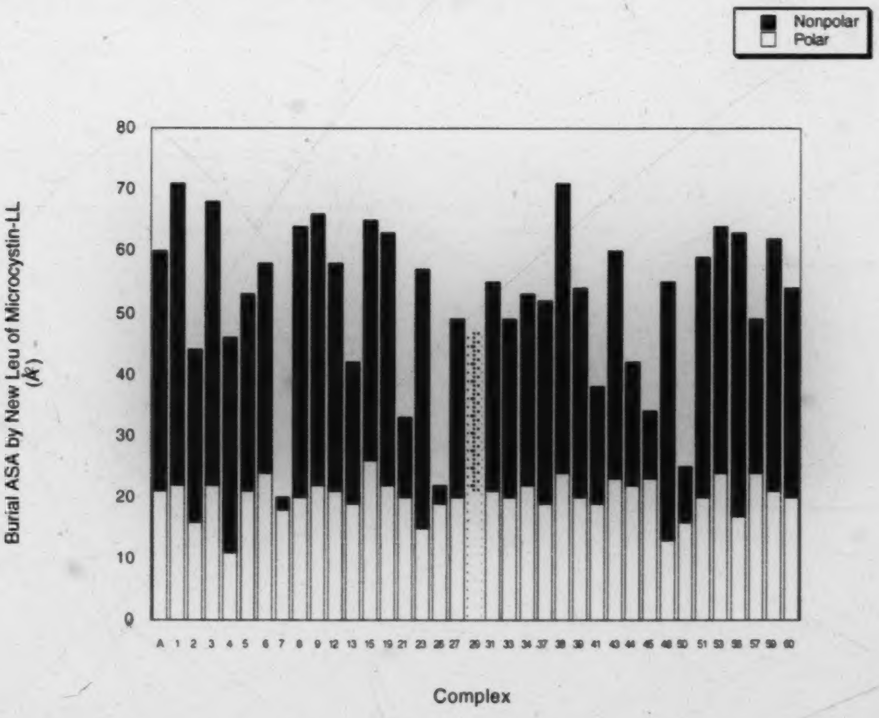


Figure 49- *Burial of the Leu residue of MLL that replaced the Arg of MLR for the high affinity MLL:PP-1c complexes. The average burial of nonpolar accessible surface area by the new Leu is 32 Å^2 (standard deviation = 12 Å^2, standard error = 2 Å^2), while burial of polar area is 21 Å^2 (standard deviation = 3 Å^2, standard error= 1 Å^2).*

Figure 50 is the burial of accessible surface area for the second Leu of MLL that is also conserved in MLR for the high affinity MLL:PP-1c complexes. The average burial of nonpolar accessible surface area by the conserved Leu is 40 Å^2 (standard deviation = 37 Å^2, standard error = 6 Å^2), while the polar burial is 0 Å^2 (standard deviation = 1 Å^2, standard error = 0 Å^2). In comparison to the high affinity MLR:PP-1c complexes (Figure 25, average nonpolar burial of Leu is 66 Å^2, standard error = 3 Å^2, while the average polar burial is 1 Å^2, standard error =0 Å^2) there is a significant reduction in the average nonpolar accessible surface area buried for the high affinity MLL:PP-1c complexes by approximately 20 Å^2 based on standard error. Also, for the MLL:PP-1c complexes there is greater differences between the individual complexes in nonpolar burial with a significant minority having no or little burial of this type of area. Thus, there appears to be some differences in the docked conformations for this Leu when comparing MLL:PP-1c verses MLR:PP-1c.

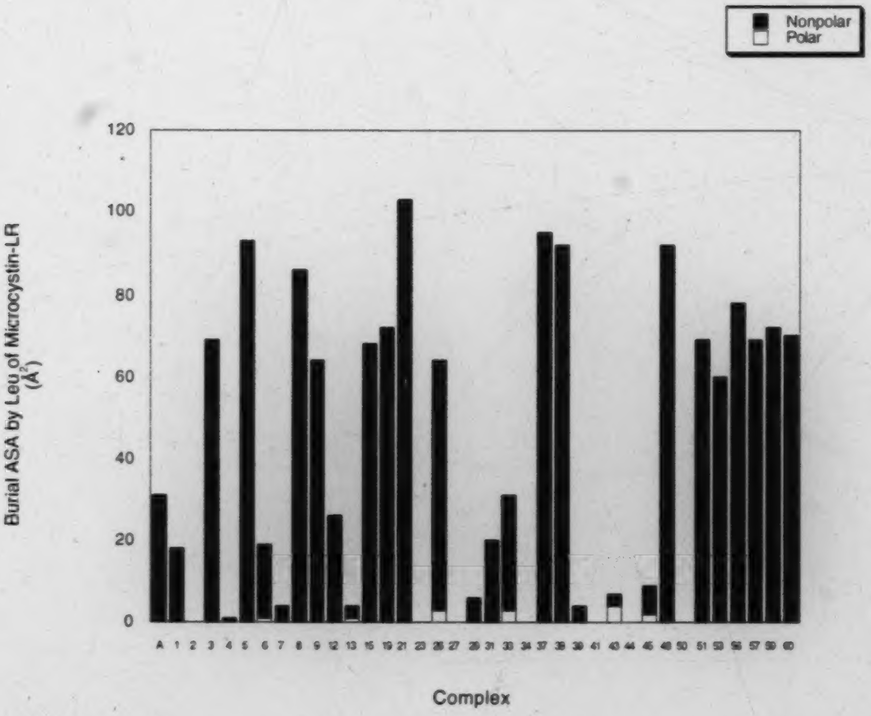


Figure 50- *Burial of accessible surface area for the second Leu of MLL that is also conserved in MLR for the high affinity MLL:PP-1c complexes. The average burial of nonpolar accessible surface area by the conserved Leu is 40 Å^2 (standard deviation = 37 Å^2, standard error = 6 Å^2), while the polar burial is 0 Å^2 (standard deviation = 1 Å^2, standard error = 0 Å^2).*

Figure 51 shows the burial of accessible surface area for the backbone residues (Ala, Masp, Glu, and Mdha) for the MLL:PP-1c high affinity complexes. The average nonpolar burial for Ala is 29 Å^2 (standard deviation = 15 Å^2, standard error = 2 Å^2), while the average polar burial is 4 Å^2 (standard deviation = 4 Å^2, standard error = 1 Å^2). The average nonpolar burial for Masp is 7 Å^2 (standard deviation = 4 Å^2, standard error = 1 Å^2), while the average polar burial is 29 Å^2 (standard deviation = 7 Å^2, standard error = 1 Å^2). For Glu, the average nonpolar burial is 26 Å^2 (standard deviation = 9 Å^2, standard error = 1 Å^2), while the average polar burial is 56 Å^2 (standard deviation = 16 Å^2, standard error = 3 Å^2). For Mdha, the average nonpolar burial is 68 Å^2 (standard deviation = 20 Å^2, standard error = 3 Å^2), while the average polar burial is 17 Å^2 (standard deviation = 9 Å^2, standard error = 1 Å^2). Similar to the MLR:PP-1c high affinity complexes, the total area for the 4 backbone residues in the MLL:PP-1c high affinity complexes were normally near or above 200 Å^2.

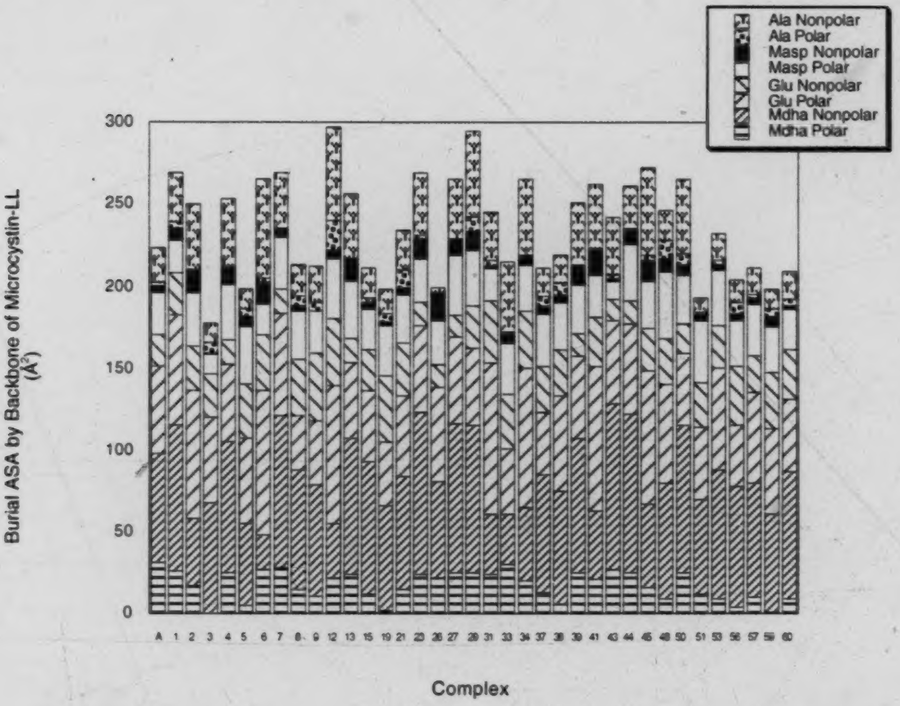


Figure 51- *Burial of accessible surface area for the backbone residues (Ala, Masp, Glu, and Mdha) for the MLL:PP-1c high affinity complexes. The average nonpolar burial for Ala is 29 Å² (standard deviation = 15 Å², standard error = 2 Å²), while the average polar burial is 4 Å² (standard deviation = 4 Å², standard error = 1 Å²). The average nonpolar burial for Masp is 7 Å² (standard deviation = 4 Å², standard error = 1 Å²), while the average polar burial is 29 Å² (standard deviation = 7 Å², standard error = 1 Å²). For Glu, the average nonpolar burial is 26 Å² (standard deviation = 9 Å², standard error = 1 Å²), while the average polar burial is 56 Å² (standard deviation = 16 Å², standard error = 3 Å²). For Mdha, the average nonpolar burial is 68 Å² (standard deviation = 20 Å², standard error = 3 Å²), while the average polar burial is 17 Å² (standard deviation = 9 Å², standard error = 1 Å²).*

Analysis of NMR docked MLL:PP-1c complexes:Part 3-20 Low Affinity Complexes

Figure 52 displays the 20 low affinity MLL:PP-1c complexes have a free energy of dissociation less than 12 kcal/mol. There is a large degree in backbone and side chain structural variation within the 20 low affinity MLL:PP-1c complexes. As with the low affinity MLR:PP-1c complexes there is almost no probability of these complexes potentially existing as observed in Figure 52. The orientations lack the a sufficient burial of nonpolar and/or polar accessible surface areas to account for the tight binding observed in MLL:PP-1c complexes.

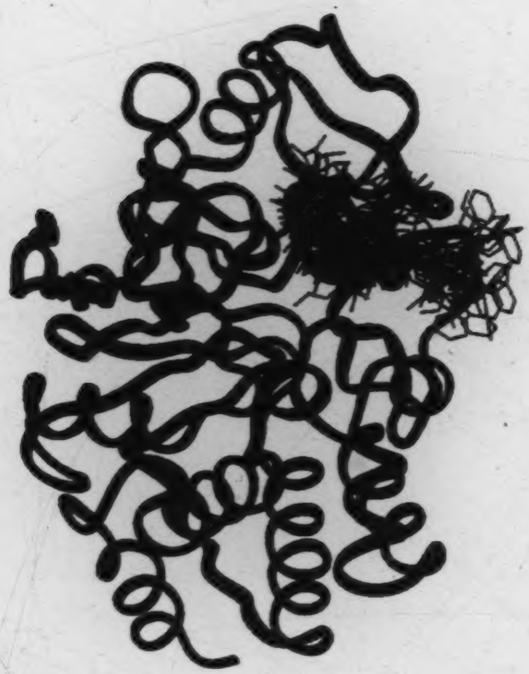


Figure 52- The 20 low affinity MLL:PP-1c complexes have a free energy of dissociation less than 12 kcal/mol. There is a large degree in backbone and side chain structural variation within the 20 low affinity MLL:PP-1c complexes.

Figure 53 is the free energy of dissociation for the 20 low affinity MLL:PP-1c complexes displayed in Figure 52. The average free energy of dissociation is 10.6 kcal/mol (standard deviation = 1.5 kcal/mol, standard error = 0.3 kcal/mol). This average is over 1 kcal/mol lower than the average for every MLL:PP-1c complex and is over 2 kcal/mol lower than the average for the high affinity MLL:PP-1c complexes.

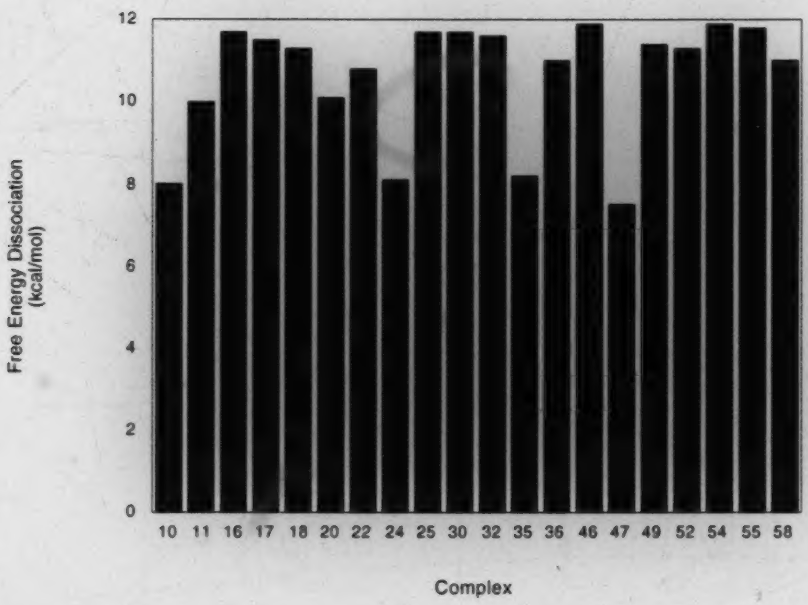


***Figure 53-** Free energy of dissociation for the 20 low affinity MLL:PP-1c complexes displayed in Figure 52. The average free energy of dissociation is 10.6 kcal/mol (standard deviation = 1.5 kcal/mol, standard error = 0.3 kcal/mol).*

Figure 54 shows the dissociation entropy and enthalpy for the 20 low affinity MLL:PP-1c complexes. The average dissociation entropy is 10.4 kcal/K/mol (standard deviation = 2.1 kcal/K/mol, standard error = 0.5 kcal/K/mol), while the average dissociation enthalpy is 0.2 kcal/mol (standard deviation = 2 kcal/mol, standard error = 0.4 kcal/mol). Of interest is the general lack of favorable binding enthalpy for this group of complexes. In fact, many of the complexes have unfavorable binding enthalpy as observed as negative dissociation enthalpy. Thus, binding by these complexes is generally driven by entropy with little or hindering contributions by enthalpy. For this set of complexes its almost like the entropy-enthalpy compensation isn't occuring, but instead there is an entropy-enthalpy battle, where the two forces are directly opposing each other. This probably is due to the lack of polar burial provided by Arg in MLR but does not occur for MLL.

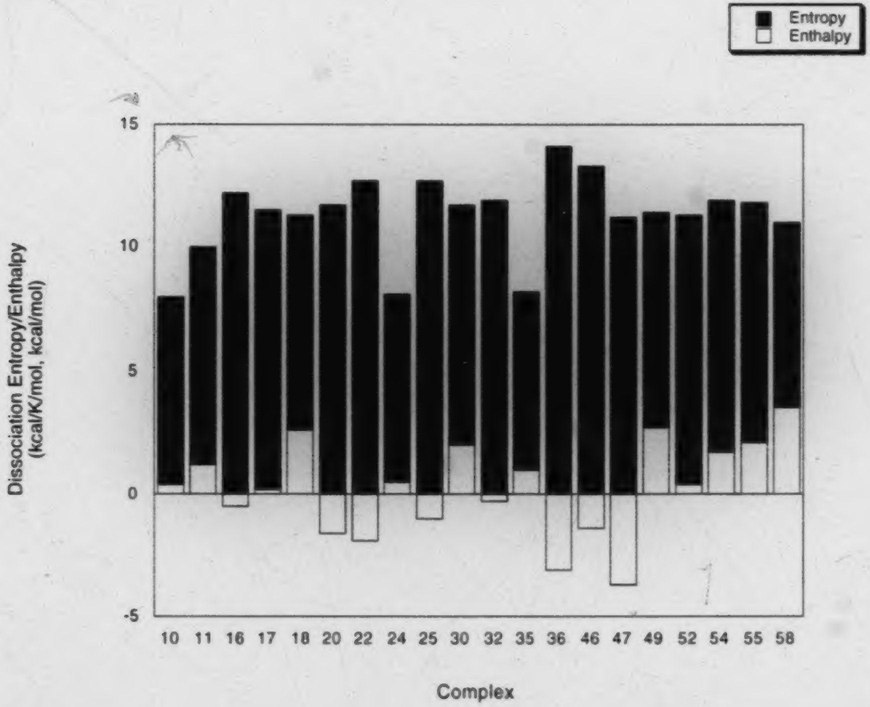


***Figure 54-** Dissociation entropy and enthalpy for the 20 low affinity MLL:PP-1c complexes. The average dissociation entropy is 10.4 kcal/K/mol (standard deviation = 2.1 kcal/K/mol, standard error = 0.5 kcal/K/mol), while the average dissociation enthalpy is 0.2 kcal/mol (standard deviation = 2 kcal/mol, standard error = 0.4 kcal/mol).*

Figure 55 examines the burial of nonpolar and polar accessible surface area by MLL in the 20 MLL:PP-1c low affinity complexes. The average nonpolar burial of accessible surface area by MLL is 361 Å^{2} (standard deviation = 48 Å^{2}, standard error = 11 Å^{2}), while the average polar burial is 126 Å^{2} (standard deviation = 26 Å^{2}, standard error = 6 Å^{2}). In comparison to the 37 high affinity MLL:PP-1c complexes (Figure 46) based on standard error, the 20 low affinity MLL:PP-1c complexes bury signficantly less nonpolar and polar accessible surface area by approximately 40 Å^{2} and 20 Å^{2} respectively. There still is though an almost 3 fold difference in the burial of these types of area with nonpolar significantly being more dominant than polar. Thus, one major reason why the MLL:PP-1c low affinity complexes do not bind as tightly is a reduction in burial of MLL.

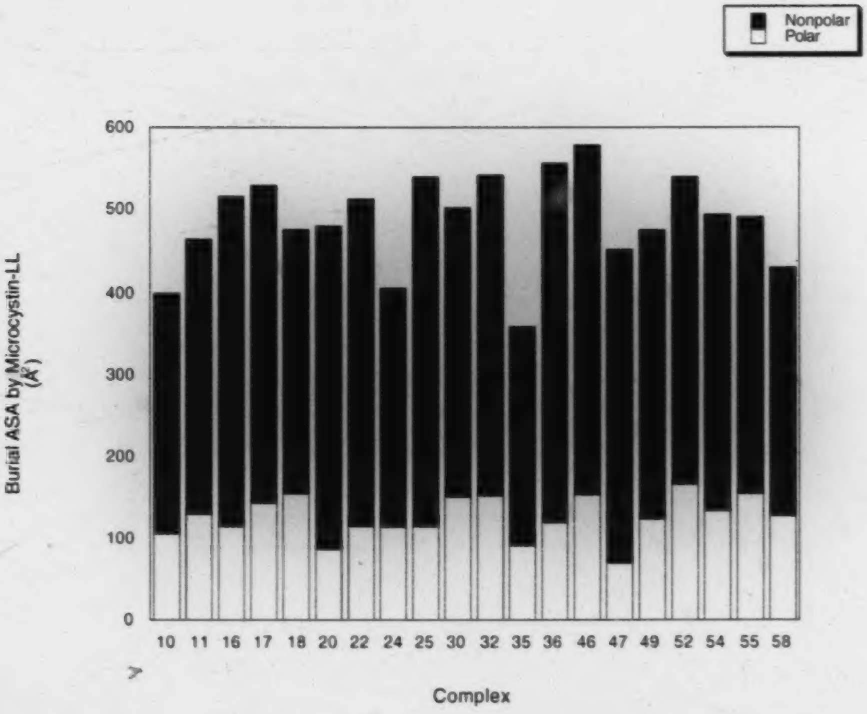


Figure 55- *Burial of nonpolar and polar accessible surface area by MLL in the 20 MLL:PP-1c low affinity complexes. The average nonpolar burial of accessible surface area by MLL is 361 Å^{2} (standard deviation = 48 Å^{2}, standard error = 11 Å^{2}), while the average polar burial is 126 Å^{2} (standard deviation = 26 Å^{2}, standard error = 6 Å^{2}).*

The other reason for why low affinity MLL:PP-1c complexes do not have tight affinity is provided in Figure 56 which displays the burial of nonpolar and polar accessible surface area by PP-1c in the 20 MLL:PP-1c low affinity complexes. The average nonpolar burial of accessible surface area by PP-1c is 216 Å^2 (standard deviation = 29 Å^2, standard error = 6 Å^2), while the average polar burial is 224 Å^2 (standard deviation = 25 Å^2, standard error = 6 Å^2). As with the MLL component of the complex, when comparing the PP-1c burial to the 37 high affinity MLL:PP-1c complexes (Figure 47), the 20 low affinity MLL:PP-1 complexes have significantly less nonpolar and polar burial of accessible surface area based on standard error by approximately 10 Å^2 and 30 Å^2 respectively. The magnitude of the differences are slightly less than the MLL component mentioned above.

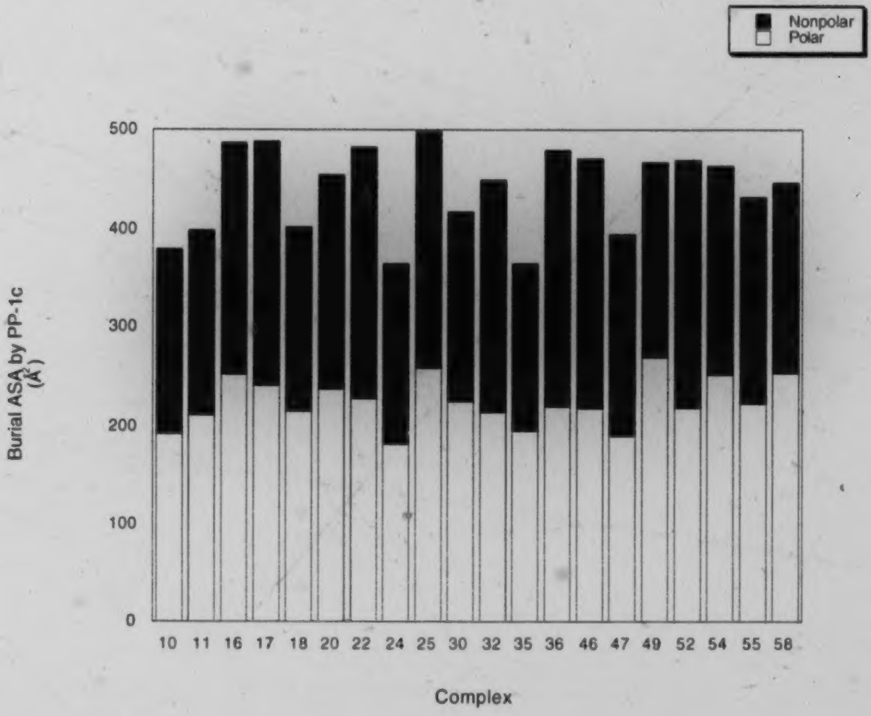


Figure 56- *Burial of nonpolar and polar accessible surface area by PP-1c in the 20 MLL:PP-1c low affinity complexes. The average nonpolar burial of accessible surface area by PP-1c is 216 Å^2 (standard deviation = 29 Å^2, standard error = 6 Å^2), while the average polar burial is 224 Å^2 (standard deviation = 25 Å^2, standard error = 6 Å^2).*

Figures 57-60 look at the specific areas of MLL that would explain the reduction in burial of accessible surface area for the 20 low affinity complexes. Figure 57 focuses on the burial of nonpolar and polar accessible surface area by Adda of MLL for the 20 MLL:PP-1c low affinity complexes. The average burial of nonpolar accessible surface area by Adda is 209 Å^2 (standard deviation = 45 Å^2, standard error = 10 Å^2), while the average burial of polar area is 23 Å^2 (standard deviation = 5 Å^2, standard error = 1 Å^2). Compared to the 37 MLL: PP-1c high affinity complexes (Figure 48), the 20 MLL:PP-1c low affinity complexes have on average no significant difference in burial of nonpolar area and the significant difference of 1 Å^2 for the polar area is so trivial that it makes the difference irrelevant. Looking at the complexes on an individual basis in Figure 57, complexes #18, #35, and #54 have a total burial of the Adda side chain well below 200 Å^2 which may account for why these complexes are of low affinity.

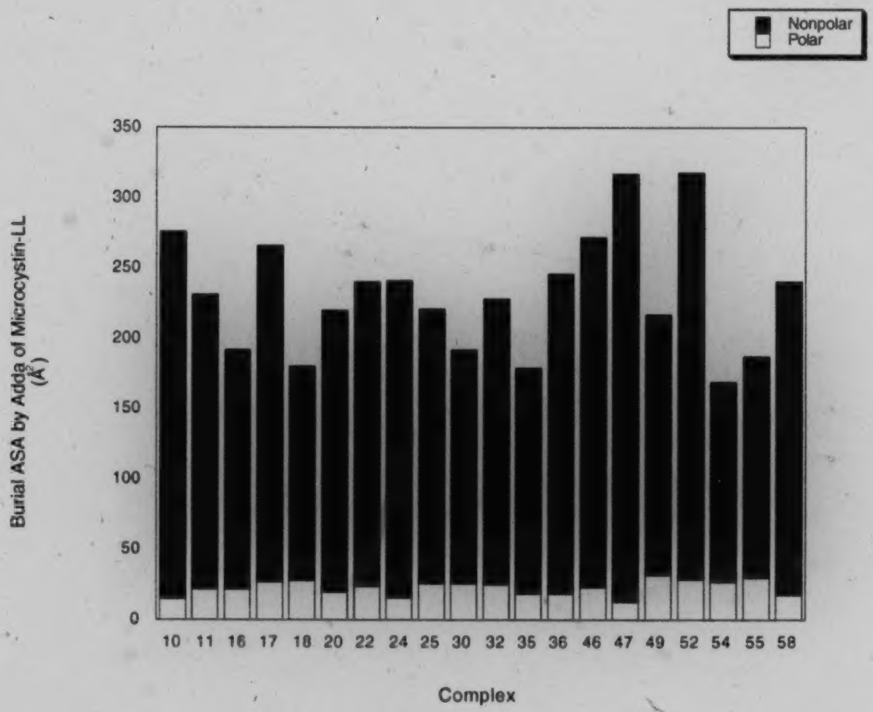


Figure 57- *Burial of nonpolar and polar accessible surface area by Adda of MLL for the 20 MLL:PP-1c low affinity complexes. The average burial of nonpolar accessible surface area by Adda is 209 Å^2 (standard deviation = 45 Å^2, standard error = 10 Å^2), while the average burial of polar area is 23 Å^2 (standard deviation = 5 Å^2, standard error = 1 Å^2).*

Figure 58 examines the burial of nonpolar and polar accessible surface area for the new Leu of MLL (replacing Arg in MLR) for the 20 low affinity MLL:PP-1c complexes. The average nonpolar burial of accessible surface area by the new Leu of MLL is 26 Å^2 (standard deviation = 17 Å^2, standard error = 4 Å^2), while the average polar burial is 14 Å^2 (standard deviation = 8 Å^2, standard error = 2 Å^2). There is a high degree of variability in the burial patterns of accessible surface area for this residue in the low affinity complexes. Some even have no contribution from this Leu towards binding due to no burial of area at all. Like Adda, when comparing the low affinity with high affinity (Figure 49) MLL:PP-1c complexes there is no significant difference on average for nonpolar area and the difference in polar area is so trivial that it is not relevant. Individually, complexes #10, #24, #35, #47, and #52 have no or very minor burial of this Leu which may account in whole or in part for their lack of tight binding.

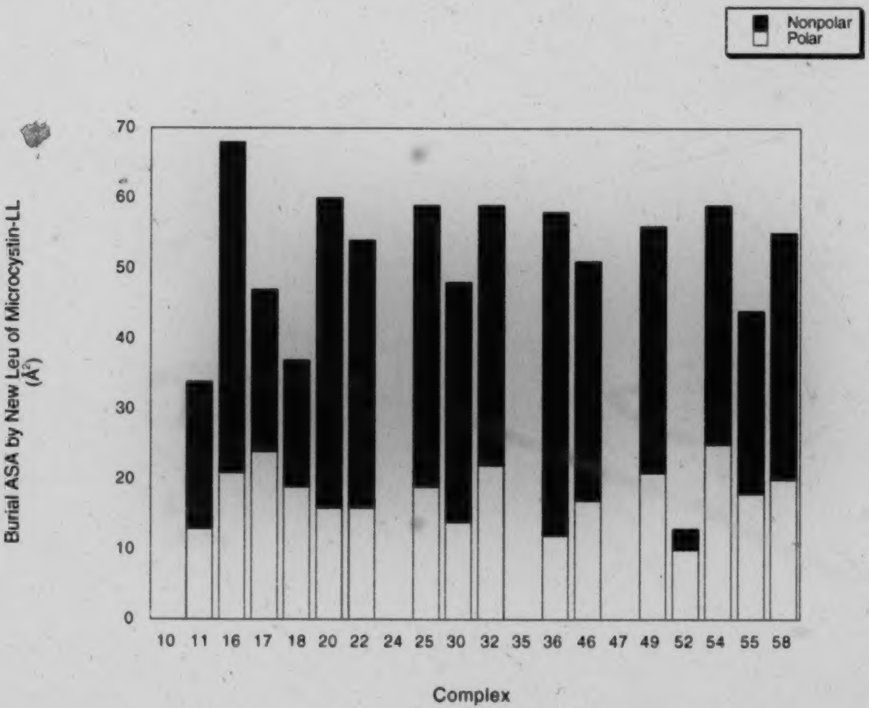


Figure 58- *Burial of nonpolar and polar accessible surface area for the new Leu of MLL (replacing Arg in MLR) for the 20 low affinity MLL:PP-1c complexes. The average nonpolar burial of accessible surface area by the new Leu of MLL is 26* $Å^2$ *(standard deviation = 17* $Å^2$*, standard error = 4* $Å^2$*), while the average polar burial is 14* $Å^2$ *(standard deviation = 8* $Å^2$*, standard error = 2* $Å^2$*).*

Figure 59 shows the burial of nonpolar and polar accessible surface area for the other Leu which is conserved between MLR and MLL for the 20 low affinity MLL:PP-1c complexes. The average burial of nonpolar accessible surface area by the conserved Leu of MLL is 23 Å^2 (standard deviation = 29 Å^2, standard error = 6 Å^2), while the average polar burial is 0 Å^2 (standard deviation = 0 Å^2, standard error = 0 Å^2). There is a trivial significant difference between the low affinity and high affinity (Figure 50) MLL:PP-1c complexes in the burial of nonpolar area while no difference exists for polar area on average based on standard error. Unlike the previous Leu where on an individual basis complexes with little or no burial of the residue could account for why the complexes where low affinity the same type of judgement can not be made for this Leu because some of the high affinity complexes had little or no burial of this residue (Figure 50).

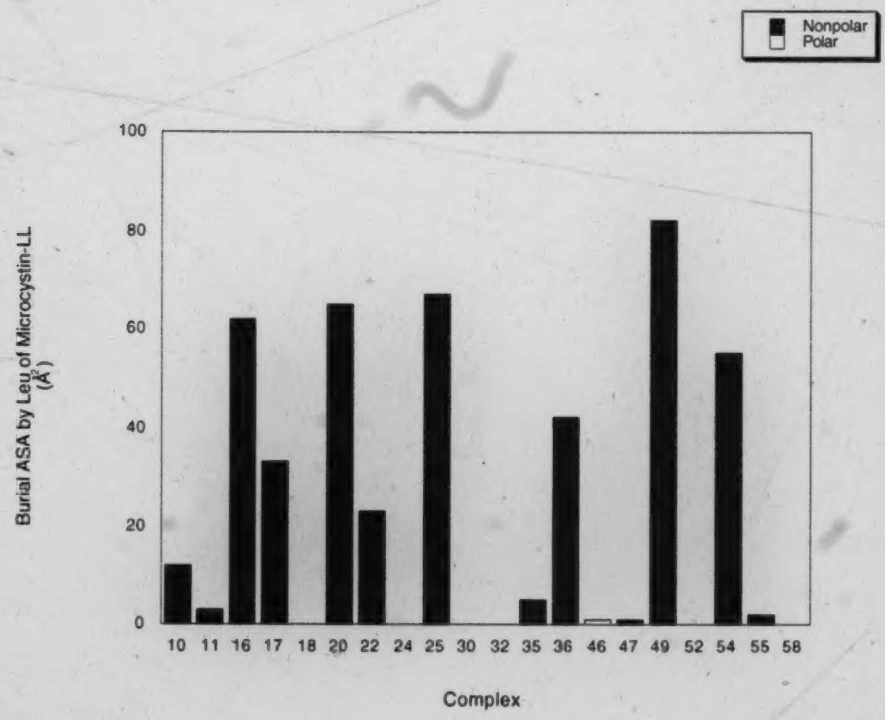


Figure 59- *Burial of nonpolar and polar accessible surface area for the other Leu which is conserved between MLR and MLL for the 20 low affinity MLL:PP-1c complexes. The average burial of nonpolar accessible surface area by the conserved Leu of MLL is 23 Å^2 (standard deviation = 29 Å^2, standard error = 6 Å^2), while the average polar burial is 0 Å^2 (standard deviation = 0 Å^2, standard error = 0 Å^2).*

Figure 60 displays the burial of nonpolar and polar accessible surface area by the backbone residues of MLL (Ala, Masp, Glu, and Mdha) for the 20 low affinity MLL:PP-1c complexes. The average nonpolar burial for Ala is 18 Å^2 (standard deviation = 14 Å^2, standard error = 3 Å^2), while the average polar burial is 2 Å^2 (standard deviation = 3 Å^2, standard error = 1 Å^2). The average nonpolar burial for Masp is 6 Å^2 (standard deviation = 4 Å^2, standard error = 1 Å^2), while the average polar burial is 24 Å^2 (standard deviation = 9 Å^2, standard error = 2 Å^2). For Glu, the average nonpolar burial is 19 Å^2 (standard deviation = 9 Å^2, standard error = 2 Å^2), while the average polar burial is 47 Å^2 (standard deviation = 10 Å^2, standard error = 2 Å^2). For Mdha, the average nonpolar burial is 61 Å^2 (standard deviation = 29 Å^2, standard error = 6 Å^2), while the average polar burial is 16 Å^2 (standard deviation = 10 Å^2, standard error = 2 Å^2). The 37 high affinity MLL:PP-1c complexes had total accessible surface area burial for the backbone residues at least near 200 Å^2 (Figure 51). Complexes #10, #20, #47, #49, and #58 have a total burial well below 200 Å^2 indicating that one reason for why these complexes are low affinity is the reduction in burial of the backbone of MLL.

In summary, MLL:PP-1c complexes #10 and #47 had low affinity because of a lack of burial of the new Leu and backbone residues. Complex #35 had reduced Adda burial and poor new Leu burial accounting for its reduction in affinity. Complexes #18 and #54 only had poor Adda residue burial to account for its low affinity. Complexes #24 and #52 lacked proper burial of the new Leu residue. Complexes #20, #49, and #58 all had poor backbone residue burial. The remaining complexes (#11, #16, #17, #18, #22, #25, #30, #32, #36, #46, and #55) had no apparent specific reason for why they were of low affinity other than on average having less MLL and PP-1c burial.

Table 2 is an overall view of the STC results from pages 112-156 for the MLR:PP-1c and MLL:PP-1c complexes.

Table 2- STC results for MLR:PP-1c and MLL:PP-1c complexes

Average Burial ASA	MLR:PP-1c High Affinity	MLR:PP-1c Low Affinity	MLL:PP-1c High Affinity	MLL:PP-1c Low Affinity
MLR Nonpolar	420 (49,9)	365 (38,10)	420 (61,10)	361 (48,11)
MLR Polar	185 (22,4)	136 (30,8)	154 (22,4)	126 (26,6)
PP-1c Nonpolar	257 (27,5)	219 (26,7)	238 (24,4)	216 (29,6)
PP-1c Polar	292 (20,4)	244 (23,6)	263 (25,4)	224 (25,6)
Adda Nonpolar	216 (47,25)	214 (52,14)	219 (53,9)	209 (45,10)
Adda Polar	25 (3,1)	20 (7,2)	26 (3,1)	23 (5,1)
MLR Arg Nonpolar	29 (12,2)	28 (14,4)	-	-
MLR Arg Polar	48 (20,4)	25 (10,3)	-	-
MLL Leu Nonpolar	-	-	32 (12,2)	26 (17,4)
MLL Leu Polar	-	-	21 (3,1)	14 (8,2)
Leu Nonpolar	66 (16,3)	34 (16,4)	40 (37,6)	23 (29,6)
Leu Polar	1 (1,0)	0 (1,0)	0 (1,0)	0 (0,0)
Ala Nonpolar	17 (8,2)	4 (6,4)	29 (15,2)	18 (14,3)
Ala Polar	3 (3,1)	2 (3,1)	4 (4,1)	2 (3,1)
Masp Nonpolar	4 (1,0)	4 (2,1)	7 (4,1)	6 (4,1)
Masp Polar	24 (4,1)	24 (7,2)	29 (7,1)	24 (9,2)
Glu Nonpolar	24 (7,1)	24 (7,2)	26 (9,1)	19 (9,2)
Glu Polar	63 (7,1)	58 (12,3)	56 (16,3)	47 (10,2)
Mdha Nonpolar	65 (14,2)	57 (26,7)	68 (20,3)	61 (29,6)
Mdha Polar	21 (11,2)	7 (9,2)	17 (9,1)	16 (10,2)

All values in Å². Numbers in brackets are standard deviation and error respectively.

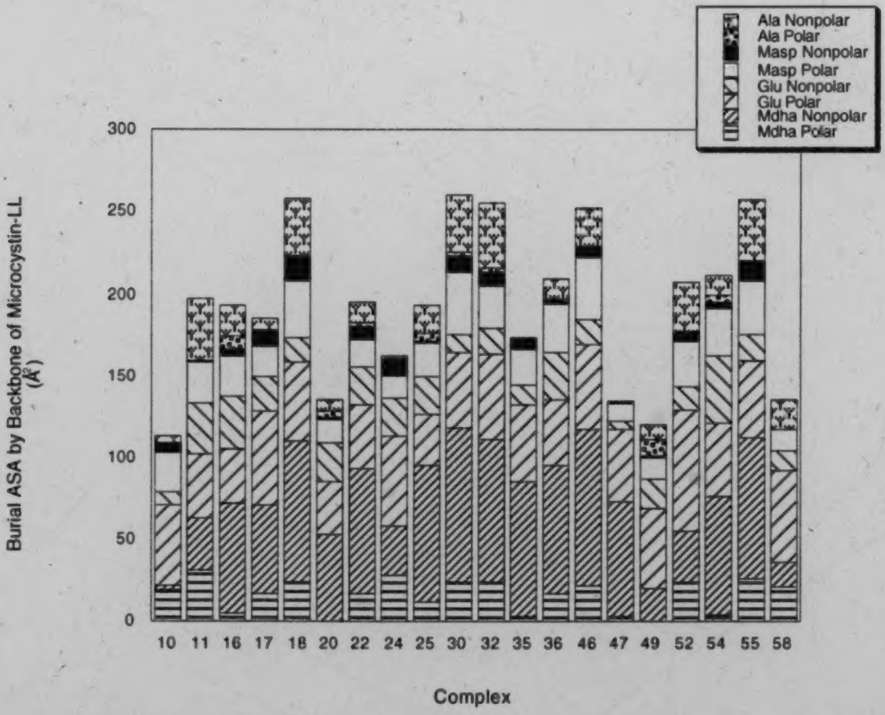


***Figure 60-** Burial of nonpolar and polar accessible surface area by the backbone residues of MLL (Ala, Masp, Glu, and Mdha) for the 20 low affinity MLL:PP-1c complexes. See text for values.*

Analysis of NMR docked MLL:PP-1c complexes:Part 4-6 Best Complexes

Figure 61 shows the docked complexes #6, #12, #21, #38, #44, and #48 for MLL:PP-1c. These complexes had the highest free energy of dissociations (Figure 44) having values no lower than 14 kcal/mol. On this basis they represent the best docked MLL:PP-1c complexes. Interestingly, they have similar backbone and Adda side chain orientations to the most probable MLR:PP-1c complexes (Figures 9 and 13 (*right*)). With respect to the Adda side chain, MLL:PP-1c complex #44 has very similar structure and position to the MLR:PP-1c crystal complex. The remaining best MLL:PP-1c complexes have Adda side chains similar to the other major orientation described in Part 4 of the MLR:PP-1c analysis. Therefore, the 6 best MLL:PP-1c complexes have strong similarities to the 7 most probable MLR:PP-1c complexes in the manner in which they bind. This may account for why there is little difference in experimental binding affinity measured for both MLR and MLL because the most probable structures for MLR and the best affinity structures for MLL are highly similar.

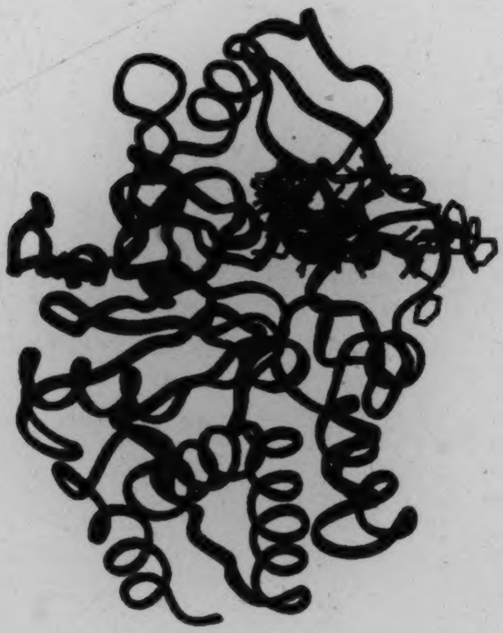


***Figure 61-** The docked complexes #6, #12, #21, #38, #44 (blue), and #48 for MLL:PP-1c. These complexes had the highest free energy of dissociations (Figure 44) having values no lower than 14 kcal/mol.*

References

1. Cohen P. 1989. The structure and regulation of protein phosphatases. *Annu Rev Biochem 58*:453-508.

2. Shenolikar S. (1994). Protein serine/threonine phosphatases-new avenues for cell regulation. *Annu Rev Cell Biol 10*:55-86.

3. Barford D. 1995. Protein phosphatases. *Curr Opin Struct Biol 5*:728-734.

4. Ohta T, Nishiwaki R, Yatsunami J, Komori A, Suganuma M, Fujuki H. 1992. Hyperphosphorylation of cytokeratins 8 and 18 by microcystin-LR, a new liver tumor promoter, in primary cultured rat hepatocytes. *Carcinogenesis 13*:2443-2447.

5. Craig M, McCready RL, Luu HA, Smillie MA, Dubord P, Holmes CFB. 1993. Identification and characterization of hydrophobic microcystins in canadian freshwater cyanobacteria. *Toxicon 31*:1541-9.

6. Mackintosh RW, Dalby KN, Campbell DG, Cohen PTW, Cohen P, and Mackintosh C. 1995. *FEBS Lett. 371*:236-240.

7. Runnegar M, Berndt N, Kong SM, Lee EYC, Zhang L. 1995. *Biochem. Biophys. Res. Commun. 216*:162-169.

8. Bagu JR, Sönnichsen FD, Williams D, Andersen RJ, Sykes BD, Holmes, CFB. 1995. Comparison of the solution structures of microcystin-LR and motuporin. *Nature Struct Biol 2*: 114-116.

9. Goldberg J, Huang H-b, Kwon Y-b, Greengard P, Nairn AC, Kuriyan, J. 1995. Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. *Nature 376*: 745-753.

10. Bagu JR, Sykes BD, Craig MM, Holmes CFB. 1997. A molecular basis for different interactions of marine toxins with protein phosphatase-1. *J Biol Chem 272*. 5087-5097.

11. Cummings MD, Hart TN, Read RJ. 1995. Atomic solvation parameters in the analysis of protein-protein docking results. Protein Sci 4: 2087-2099.

12. Vajda S, Sippl M, Novotny J. 1997. Empirical potentials and functions for protein folding and binding. *Curr Opin Struct Biol 7*: 222-228.

13. Abagyan, R. & Totrov, M. Biased probability Monte Carlo conformational searches and electrostatic calculations for peptide and proteins. *J Mol Biol* **235**, 983-1002 (1994).

14. Jackson RM, Sternberg MJE. 1995. A continuum model for protein-protein interactions: Application to the docking problem. *J Mol Biol 250*: 258-275.

15. Gilson MK, Given AG, Bush BL, McCammon JA. 1997. The statistical thermodynamic basis for computation of binding affinities: A critical review. *Biophys J* 72: 1047-1069.

16. Weng Z, Vajda D, DeLesi C. 1996. Prediction of complexes using empirical free energy functions. *Protein Sci 5*:614-626.

17. Novotny J, Bruccoleri RE, Davis M, Sharp KA. 1997. Empirical free energy calculations: A blind test and further improvement of the method. *J Mol Biol 268*:401-411.

18. Murphy KP, Freire E. 1992. Thermodynamics of structural stability and cooperative folding behaviour in proteins. *Adv Protein Chem 43:* 313-361.

19. Luque I, Freire E. 1998. Structure-based prediction of binding affinities and molecular design of peptide ligands. *Methods Enzymol 295*: 295-315.

20. Baker BM, Murphy, KP. 1998. Prediction of binding energetics from structure using empirical parameterization. *Methods Enzymol 295*: 295-315.

21. Baker BM, Murphy KP. 1997. Dissecting the energetics of a protein-protein interaction: The binding of ovomucoid third domain to elastase. *J Mol Biol 268*: 557-569.

22. Dauber-Ogusthorpe P, Roberts VA, Ogusthorpe DJ, Wolff J, Genest M, Hagler AT. 1988. Structure and energetics of ligand binding to proteins: *E. coli* dihydrofolate reductase-trimethoprim, a drug-receptor system. *Proteins Struct Funct Genet 4*: 31-47.

23. Gomez J, Freire E. 1996. Thermodynamic mapping of the inhibitor site of the aspartic protease endopepstatine. *J Mol Biol 252*:337-350.

24. Xie D, Freire E. 1994. Molecular basis of cooperativity in protein folding. V. Thermodynamic and structural conditions for the stabilization of compact denatured states. *Proteins Struct Funct Genet 19*:291-301.

25. Hilser VJ, Gomez J, Freire E. 1996. The enthalpy change in protein folding and binding: refinement of parameters for structure -based calculations. *Proteins Struct Funct Genet 26*:123-133.

26. Baldwin RL. 1986. Temperature dependence of the hydrophobic interaction in protein folding. *Proc Natl Acad Sci USA 83*, 8069-8072.

27. D'Aquino JA, Gomez J, Hilser VJ, Lee KH, Amzel LM, Freire E. 1996. The magnitude of the backbone conformational entropy change in protein folding. *Proteins Struct Funct Genet 25:*143-156.

28. Murphy KP, Xie D, Thompson KS, Amzel LM, Freire E. 1994. Entropy in biological processes: Estimation of the translational entropy loss. *Proteins Struct Funct Genet 18* :62-67.

29. Lee KH, Xie D, Freire E, Amzel LM 1994. Estimation of changes in side-chain configurational entropy in binding and folding: General methods and application to helix formation. *Proteins Struct Funct Genet 20* :68-84.

30. Egloff MP, Cohen PTW, Reinemer P, Barford, D. 1995. Crystal structure of the catalytic subunit of human protein phosphatase 1 and its complex with tungstate. *J Mol Biol 254:*942-959.

31. Miller S. Janin J, Lesk AM, Chothia C. 1987. Interior and surface of proteins. *J Mol Biol 196*: 641-656.

32. Bardi J, Luque I, Freire E. 1997. Structure-based thermodynamic analysis of HIV-1 protease inhibitor. *Biochemistry 36*: 6588-6596.

33. Pickett S, Sternberg MJE. 1993. Empirical scale of side-chain conformational entropy in protein folding. *J Mol Biol 231*: 825-839.

34. Kauzmann W. 1959. Some factors in the interpretation of protein denaturation. *Adv Protein Chem 14*: 1-63.

35. Holtzer A. 1995. The "cratic correction" and related fallacies. *Biopolymers 35*: 595-602.

36. Tamura A, Privalov PL. 1997. The entropy of protein association. J Mol Biol 273: 1048-1060.

37. Yu Y, Lavigne P, Kay CM, Hodges RS, Privalov PL. 1998. Contribution of translational and rotational entropy to the unfolding of a dimeric coiled-coil. *J Phys Chem B 103*: 2270-2278.

38. Amzel LM. 1996. Loss of translational entropy in binding, folding and catalysis. *Proteins Struct Funct Genet 29:* 1-6.

39. Richmond TJ. 1984. Solvent accessible surface area and excluded volumes in proteins. *J Mol Biol 82*:63-89.

40. Wishart DS, Willard L, Richards FM, Sykes BD. 1994. VADAR: A comprehensive program for protein structure evaluation. Version 1.2, University of Alberta, Edmonton, Alberta, Canada.

41. Nicholls A, Sharp KA, Honig B. 1991. Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct Funct Genet 11*:281-296.

42. Nishiwaki-Matsushima R, Nishiwaki S, Ohta T, Yoshizawa S, Suganuma M, Harada K-I, Watanabe MF, Fujiki H. 1991. Structure-function relationships of microcystin, liver tumor promoters, in interaction with protein phosphatase. *Jpn J Cancer Res 82*: 993-996.

43. Stotts RR, Namikoshi M, Haschek WM, Rinehart KL, Carmichael WW, Dahlem AM, Beasley VR. 1993. Structural modifications imparting toxicity in microcystins from microcystis spp. *Toxicon 41*: 783-789.

44. Huang H-B, Horiuchi A, Goldberg J, Greengard P, Nairn AC. 1997. Site-directed mutagenesis of amino acid residues of protein phosphatase 1 involved in catalysis and inhibitor binding. *Proc Natl Acad Sci USA 94*: 3530-3535.

45. Bradi GP, Sharp KA. 1997. Entropy in protein folding and in protein-protein interactions. *Curr Opin Struct Biol 7*: 215-221.

46. Takai A, Sasaki K, Nagai H, Mieskes G, Isobe M, Isono K, Yasumoto T. 1995. Inhibition of specific binding of okadaic acid to protein phosphatase-2A by microcystin-LR, calyculin-A and tautomycin: method of analysis of interactions of tight-binding ligands with target protein. *Biochem J 306*: 657-665.

47. Lumry R, Rajender S. 1970. Enthalpy-entropy compensation phenomena in water solutions of proteins and small molecules: A ubiquitous property of water. *Biopolymers 9*: 1125-1227.

48. Lumry R. 1995. The new paradigm for protein research. In: Gregory RB, ed. *Protein-solvent interactions*. New York, Marcel Dekker, inc. pp. 1-136.

49. Hawkes R, Grutter MG, Schellman J. 1984. Thermodynamic stability and point mutations of bacteriophage T4 lysozyme. *J Mol Biol 175:* 195-212.

50. Shortle D, Meeker AK, Freire E. 1988. Stability mutants of staphyloccocal nuclease: Large compensation enthalpy-entropy changes for the reversible denaturation reaction. *Biochemistry 27*: 4761-4768.

51. Eftink MR, Anusiem AC, Biltonen RL. 1983. Enthalpy-entropy compensation and heart capacity changes for protein-ligand interactions: General thermodynamic models and data for the binding of nucleotides to ribonuclease A. *Biochemistry 22*: 3884-3896.

52. Makhatadze GI, Privalov P. 1995. Energetics of protein structure. *Adv Protein Chem 47*: 307-417.

53. Cooper A. 1976. Thermodynamic fluctuations in protein molecules. *Proc Natl Acad Sci USA 73*: 2740-2741.

54. Frauenfelder H, Parak F, Young RD. 1988. Conformational substates in proteins. *Ann Rev Biophys Chem 17*: 451-479.

Thesis Chapter 5

Structural comparison of protein phosphatase-1 with calcineurin: explanation for why microcystin is unable to inhibit calcineurin

Some figures in this chapter will be included in a paper submitted to Journal of Biological Chemistry with Bagu, J.R. as third author. Modeling, analysis of conserved residues, and other figures by Bagu, J.R..

Summary

A structural comparison of the catalytic subunits of protein phosphatase-1 (PP-1c) and protein phosphatase-2B (calcineurin) was undertaken in an attempt to explain why the marine toxin microcystin-LR inhibit PP-1c but not calcineurin. Structurally, these two enzymes are highly similar, especially in the region where microcystin binds to PP-1c. Eight residues in this region are identical in sequence and have similar three-dimensional structure. However, nine residues in the PP-1c inhibitor binding pocket are not conserved when comparing sequences. These nine residues were selectively mutated in PP-1c to the corresponding residues in calcineurin. The activities of these mutant PP-1c enzymes in the presence of an inhibitor (microcystin-LR) were then compared to native PP-1c to see if the calcineurin residues diminished binding. No single mutation significantly perturbed microcystin-LR binding. However, a major structural difference in the position of the L7 loop within the binding pocket (PP-1c residues 272-278, calcineurin residues 311-317) is postulated to account for why microcystin-LR is unable to inhibit calcineurin. The structural shift in the L7 loop in calcineurin is caused by an interaction external to the binding pocket between His 339 and Asp 313. This interaction pulls the backbone of the L7 loop in calcineurin away from the microcystin-LR binding site in PP-1c causing the side chains of two calcineurin residues, Leu 312 and Tyr 315, to potentially interfere with microcystin-LR binding.

Introduction

Protein phosphatase-1c (PP-1c) and protein phosphatase-2b (calcineurin) are two major protein serine/threonine phosphatases responsible for regulating many biochemical processes in eukaryotic cells. Their function is inherently opposite to protein kinases which are enzymes that phosphorylate a variety of different proteins. This dynamic equilibrium between phosphatases and kinases result in many processes being regulated by reversible protein phosphorylation.

PP-1c activity is normally regulated by interaction with regulatory subunits and reversible protein phosphorylation. PP-1c activity is also suppressed by a variety of marine toxins that are able to bind and inhibit this enzyme. These toxins are produced by cyanobacteria (e.g. microcystins) or dinoflagellates (e.g. okadaic acid). They are potent inhibitors of PP-1c having IC_{50} values near 0.2 nM. However, the homologous enzyme calcineurin relative to PP-1c is not inhibited by these toxins.

This functional difference is surprising considering the strong sequence homology shared by these two enzymes. PP-1c shares 40% sequence identity with calcineurin in the catalytic domain (Figure 2, Chapter 1). Table 1 is a sequential alignment of protein phosphatase-1 and calcineurin (PP-2B) using the program SEQSEE[1]. PP-1 is the human sequence, while the calcineurin sequence is bovine. The sequence comparison starts at residue 27 for calcineurin. The extended C-terminus of calcineurin is not shown in this alignment. The L7 loop previously defined[2] as important in inhibitor binding is underlined.

Table 1-SEQSEE[1] sequential alignment of PP-1c and calcineurin (PP-2B)

```
Protein 1: PROTEIN PHOSPHATASE-1
Protein 2: PROTEIN PHOSPHATASE-2B (Calcineurin)
Score....: 1819
Matches..: 114


Protein 1: MSDSEKLNLDSIIGRLLEVQGSRPGKNVQLTENEIRGLCLKSREIFLSQP   50
Matching.:      |  *|      ***  *   |   | |            |*  *
Protein 2: HRLTAKEVFDNDGKPRVDILKAHLMKEGRLEETVALRIITEGASILRQEK   76


Protein 1: ILLELEAPLKICGDIHGQYYDLLRLFEYGGFPPESNYLFLGDYVDRGKQS   100
Matching.:  ||* *||* *|||||||**||**||| || |    |||||||||||  |
Protein 2: NLLDIDAPVTVCGDIHGQFFDLMKLFEVGGSPANTRYLFLGDYVDRGYFS   126


Protein 1: LETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWK   150
Matching.:  | *  | | || ||  *|||||||||      * |  ||| *|  ***
Protein 2: IECVLYLWALKILYPKTLFLLRGNHECRHLTEYFTFKQECKIKYSERVYD   176


Protein 1: TFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQSMEQIRRIMRPTDVPDQGL   200
Matching.:    | |*||| ||  ** **| ||||||* * ** ||*  |  * | |
Protein 2: ACMDAFDCLPLAALMNQQFLCVHGGLSPEINTLDDIRKLDRFKEPPAYGP   226


Protein 1: LCDLLWSDP      DKDVQGWGEND RGVSFTFGAEVVAKFLHKHDLDL   243
Matching.: *|| |||||      *|  * *  |  || |* *    |  ||   *|
Protein 2: MCDILWSDPLEDFGNEKTQEHFTHNTVRGCSYFYSYPAVCEFLQHNNLLS   276


Protein 1: ICRAHQVVEDGYEFFAKRQ      LVTLFSAPNYCGEFDNAGAMMSVDE   287
Matching.: | |||*  * ||  * | |      |*| ||||||   **|  | *  *
Protein 2: ILRAHEAQDAGYRMYRKSQTTGFPSLITIFSAPNYLDVYNNKAAVLKYEN   326


Protein 1: TLMCSFQILKPADKNKGKYGQFSGLNPGGRPITPPRNSAKAKK          330
Matching.:  *|  *  ** *        *|  *     |*   *      *
Protein 2: NVM NIRQFNCSPHPYW LPNFMDVFTWSLPFVGEKVTEMLVNVL        369
```

The crystal structures of PP-1c:microcystin-LR[2] and the catalytic subunit of calcineurin bound to its regulatory subunit, immunosuppressant (FK506), and immunophilin (FKBP12)[3] have both been determined. This study compares PP-1c with calcineurin A to see if there are structural differences which could explain why there is a different inhibitory response to the marine toxin microcystin-LR. Mutagenesis studies of residues that differ between PP-1c and calcineurin were undertaken to determine if these residues were significant.

Methods

Molecular Modelling

Protein phosphatase-1 and calcineurin sequence alignment was performed by SEQSEE[1]. Straight lines indicated conserved residues, while stars indicate similar residues. The crystal structures of the catalytic subunits for recombinant rat PP-1c[2] and bovine calcineurin[3] were superimposed using Insight II software version 2.3 (Biosym Technologies). This superimposition resulted from superimposing only 8 structurally identical residues located in the microcystin-LR binding pocket in PP-1c (Table 3, Results and Discussion). This superimposition was chosen based on a close examination of the PP-1c residues near the microcystin-LR binding site. The 8 corresponding conserved calcineurin residues were then selected for superimpositon of the two enzymes. This type of superimposition was chosen in order to maximize the structural similarity in the microcystin-LR binding region. The resulting backbone RMSD of the two enzymes in the overlapping sequences (based on overlapping residues in Table 1, excluding calcineurin insertions located in the area from residues 226-326) was 3.33Å. Note, this is not the RMSD if one to directly superimpose the backbones of the two enzymes, it is a reflection of the perspective microcystin-LR would have in binding to the enzymes. Metropolis Monte Carlo docking of microcystin-LR individually to PP-1c[4] and to calcineurin were accomplished using the Monte Carlo macro in Insight II.

Mutagenesis and Purification of PP-1c

This work was preformed by John Dawson and Hue Anh Luu in the Charles Holmes Lab. It involved mutagenesis of PP-1c by inserting PP-1c cDNA into a sequencing vector. Mutant forms were constructed using Quikchange Mutagenesis System (Stratagene). Verification of correctness was performed on all constructs produced. Purification of recombinant mutant PP-1c was accomplished following expression in *E. coli*. Dose response curves were then determined using ^{32}P labelled phosphorylase a as a substrate for the mutant PP-1c constructs in the presence of varying conconcentration of microcystin-LR.

Results and Discussion

Comparison of the crystal structures of protein phosphatase-1c and calcineurin A

The overall architecture of MLR bound to PP-1c is shown in Figure 1[2]. MLR binds at a junction between the C-terminal groove, acidic groove, and hydrophobic groove where the Adda side chain fits. The 3 grooves form a Y-shape and are involved in the activity of PP-1c with residues in each groove used to bind the substrate. Inhibition of PP-1c by MLR occurs because bound MLR partially blocks the substrate binding site by interacting with some PP-1c residues involved in the dephosphorylation process (e.g. Arg 96 and 221).

***Figure 1-** Crystal structure complex of MLR bound to PP-1[2]. Three grooves form a Y-shape in the vicinity where MLR binds to with the Adda side chain of MLR fitting into the hydrophobic groove.*

Figure 2 shows the comparison of the crystal structure of PP-1c (without MLR)[2] with the crystal structure of calcineurin A (without calcineurin B, FKBP12, and FK506)[3] in the same orientation as Figure 1. The positions of the enzymes results from superimposing 8 conserved residues between the two enzymes in the MLR binding site (see Figure 3 and Table 3). The resulting backbone r.m.s.d. for overlapping residues (Table 1) is 3.33 Å (See Methods for explanation). There is a clear similarity in the overall architecture of PP-1c and calcineurin B. In fact, one can clearly see the 3-dimensional similarities in the catalytic sites of both phosphatases. The C-terminal groove, acidic groove, and hydrophobic groove observed in PP-1c (Figure 1) also exist in calcineurin B (Figure 2). The long C-terminal helix of calcineurin A seen at the top, left of Figure 2 binds to calcineurin B.

Figure 2- *Comparison of backbones of PP-1c (red) and calcineurin A (PP-2B in green) in ribbons*[5]*. The orientation is identical to the PP-1c orientation in Figure 1. Calcineurin A shares the same 3 grooves at its catalytic site as observed in PP-1c (Figure 1). This superimposition was based on overlapping identical residues (total of 8) in the catalytic/MLR binding site. The backbone r.m.s.d. for the comparison of PP-1c with calcineurin A is 3.33 Å. The PP-1c crystal structure has had MLR removed from it, while the calcineurin A crystal structure has had calcineurin B, FKBP12, and FK506 removed from it.*

The structural similarity observed in Figure 2 explains how both PP-1c and calcineurin A have the same Ser/Thr dephosphorylation activity. However, despite the similar tertiary structure in the catalytic sites MLR does not bind to calcineurin A while it is able to inhibit PP-1c. Analysis of the MLR binding site in PP-1c shows that 8 identical residues are shared between PP-1c and calcineurin A (Figure 3). Not only are the 8 residues are identical in a sequence comparison of the two enzymes (Table 1), but they also have highly similar 3-dimensional positions relative to one another including side chains. From the perspective of Figure 3 one might expect that MLR would be able to bind and inhibit calcineurin in the same manner as PP-1c.

Figure 3- *A comparison of 8 identical residues between PP-1c in red and calcineurin A in green. These 8 residues are within 4 Å of MLR in PP-1c and actively interact with MLR. Despite having these almost identically positioned residues, calcineurin A is not able to interact with MLR, its activity is not inhibited by MLR. These 8 residues where used as a basis for the superimposition in Figure 2.*

The similar structural motifs in the binding site of MLR in PP-1c compared to calcineurin A suggest that MLR could be docked onto calcineurin A at this location. Figure 4 shows the results of the attempted docking using the Metropolis Monte Carlo procedure in Insight II which previously was used in other dockings of toxins onto PP-1c^4. The starting position of MLR was the position in the crystal structure complex of MLR:PP-1c determined by Goldberg et al.^{2} relative to the 8 identical residues found in calcineurin A. The final position of MLR was dislocated from the binding position that MLR has in the crystal structure of the MLR:PP-1c complex. Comparison of the two positions indicates that the backbone of MLR is pushed out of the binding pocket at a range of 5-6 Å, the Arg residue is 5 Å away, while the Adda side chain is 3-4 Å away. Overall, MLR was not able to fit into the same binding position in calcineurin A that it had in PP-1c. This is reflected by the STC calculations (Appendix 1) of the final docked MLR:calcineurin complex where the dissociation constant is $4.4 * 10^{-7}$, a value greater than what was calculated for the MLR:PP-1c crystal complex ($4.1 * 10^{-11}$).

Figure 4- Attempted docking of MLR onto calcineurin A using the Metropolis Monte Carlo procedure used within Biosym's Insight II. The starting position of MLR in calcineurin is identical relative to the 8 residues shared between this enzyme and PP-1c in the exact binding position of MLR bound to PP-1c. Microcystin-LR is unable to stay in this position and maintain the same interaction that it has with PP-1c.

In order to determine why MLR is unable to dock into calcineurin A a detailed analysis of the surrounding calcineurin A residues was done using the position of MLR:PP-1c complex as a starting point. Table 2 shows all the calcineurin A residues within 4 Å of MLR shown in Figure 3. Table 2 also gives the equivalent residue to PP-1c that is found in calcineurin A based on sequence homology (Table 1).

Table 2- PP-1c and Corresponding Calcineurin Residues in the MLR Binding Site

Calcineurin A Residue	PP-1c Equivalent
Arg 122	Arg 96
Asn 150	Asn 124
His 151	His 125
His 155	Ser 129
Leu 156	Ile 130
Tyr 159	Ile 133
Phe 160	Tyr 134
Glu 220	Asp 194
Pro 221	Val 195
Pro 222	Pro 196
Cys 228	Cys 202
Trp 232	Trp 206
Val 253	Asp 220
Arg 254	Arg 221
Gly 255	Gly 222
Cys 256	Val 223
Tyr 311	Tyr 272
Leu 312	Cys 273
Val 314	Glu 275
Tyr 315	Phe 276

From these residues in Table 2 a subdivision can be made into two groups. The first group is displayed in Table 3 which are found to be identical in both PP-1c and calcineurin A. These amino acid residues are also displayed in Figure 3.

Table 3- Identical Residues in Calcineurin A and PP-1c at the MLR Binding Site

PP-1c	Calcineurin A
Arg 96	Arg 122
Asn 124	Asn 150
His 125	His 151
Pro 196	Pro 222
Trp 206	Trp 232
Arg 221	Arg 254
Gly 222	Gly 255
Tyr 272	Tyr 311

The second group are those residues that differ between PP-1c and calcineurin A at the MLR site (Table 4). The answer to why MLR is unable to inhibit calcineurin A might be expected to be found within the residues listed in Table 3.

Table 4-Non-conserved Residues When Comparing PP-2B and PP-1 at the MLR site

PP-1c	Calcineurin A
Ser 129	His 155
Ile 130	Leu 156
Ile 133	Tyr 159
Tyr 134	Phe 160
Val 195	Pro 221
Asp 220	Val 253
Val 223	Cys 256
Cys 273	Leu 312
Glu 275	Val 314
Phe 276	Tyr 315

Figure 5 shows the majority of the calcineurin A residues listed in Table 4 as a GRASP surface image. There are clear steric clashes between MLR and some of the calcineurin residues especially Leu 312, Tyr 315, and His 155. Some of these clashes might explain why MLR is unable to bind to calcineurin A and have been further investigated.

Figure 5- GRASP image of the surface of calcineurin A in the same orientation as Figure 2. The MLR position is what it would be when bound to PP-1c is exactly identical relative to the 8 shared residues in the binding pocket between PP-1c and calcineurin A. Residues that are found to be different in calcineurin A as opposed to PP-1c are labelled on the surface of calcineurin A. Red surfaces are negatively charge, blue are positive, white is neutral and yellow is hydrophobic.

Table 5 lists in point form potential reasons why MLR is unable to bind to calcineurin A based on individually comparing the pairs of residues listed in Table 4. The reasons given refer to changes in calcineurin A that would potentially block MLR binding or amino acids that lack important binding interactions between PP-1c and MLR. Some of the clashes are observable in Figure 5.

Table 5- Structural Analysis of Differing Residues Between 2 Enzymes

PP-1c	Calcineurin A	Comment
Ser 129	His 155	Longer side chain of His 155 too close to Adda.
Ile 130	Leu 156	Very Similar.
Ile 133	Tyr 159	Longer side chain of Tyr 159 might block substrate.
Tyr 134	Phe 160	Lacks hydroxyl group.
Val 195	Pro 221	Similar.
Asp 220	Val 253	Val 253 side chain too close to Arg in MLR, replacement of a negative by a neutral residue.
Val 223	Cys 256	Similar.
Cys 273	Leu 312	Leu 312 clashes with Leu in MLR.
Glu 275	Val 314	Very different positions but no conflicts with MLR.
Phe 276	Tyr 315	Different positions, Tyr 315 clash Mdha in MLR.

An important structural difference was found to be existing in the location of the catalytic site for both enzymes which has an important role in MLR binding to PP-1c. Tyr 272 of PP-1c and Tyr 311 of calcineurin A are identical in terms of side chain positions. However, the backbones start to diverge at those residues which can be seen in the phi and psi measurements of the Tyr 272/Tyr 311: ϕ,ψ = -94, -45 verses -48, +131 for PP-1c and calcineurin respectively, ψ 176 degrees different causing backbones to diverge in opposite directions. Therefore, PP-1c residues Cys 273, Gly 274, Glu 275, Phe 276, and Asp 277 have different backbone and side chain positions verses calcineurin A residues Leu 312, Asp 313, Val 314, Tyr 315, and Asn 316. Backbone and side chains return to similar positions at PP-1c Asn 278 verses Asn 317 calcineurin A. This structural deviation between PP-1c and calcineurin A is shown in Figure 6. The steric clashes that calcineurin A residues Leu 312 and Tyr 315 with Leu and Mdha of MLR are highlighted in Figure 6. This loop of residues, Tyr 311 to Asn 317, in calcineurin A is termed the L7 loop by Griffith et al.[3]. For convenience the PP-1c residues Tyr 272 to Asn 278 will also be called the L7 loop in PP-1c. The difference in backbone conformations of the L7 loop between the two enzymes result in different side chain locations for the L7 loop residues. This creates steric interference between calcineurin A and MLR which is hypothesized to disrupt binding.

Figure 6- *Stereo view of the most significant structural difference between PP-1c and calcineurin A at the catalytic sites of both enzymes*[5]*. The L7 loop of PP-1C is in red, green represents the L7 loop of calcineurin A, and blue is the bound crystal structure of MLR. Van der Waals radii for each structure is respresented in their respective colors. There is an overlap in the radii between Leu 312 of calcineurin A in green and Leu of MLR in blue and Tyr 315 of calcineurin in green and Mdha of MLR in blue. No significant overlap exists between PP-1c in red and MLR in blue. This is a consequence of the diverging backbone conformations of PP-1c and calcineurin A in the L7 loop starting at the respective Tyrs and ending at the respective Asns. The different backbone conformations results in different side chain locations in which the two previously mention sidechains of calcineurin A sterically interfere with MLR binding.*

Based on the analysis of the differing residues in Table 5 a series of individual mutations where made. These mutations were conducted on PP-1c. The goal was to try to determine which calcineurin A residues are responsible for its immunity to MLR by substituting the calcineurin A residue outlined in Table 5 in the equivalent position of PP-1c.

In point form below are the results of the mutational studies. These mutagenesis studies were mainly carried out by John Dawson of the Charles Holmes lab.

1. Ser 129 PP-1c mutated to a His. Negligible difference from native PP-1c activity in the presences of MLR when placing a His in position 129.

2. Ile 130 PP-1c, Leu 156 calcineurin A. No mutation done, both hydrophobic residues.

3. Ile 133 PP-1c mutated to a Tyr. Negligible difference from the native dose response curve in PP-1c activity.

4. Tyr 134 PP-1c switched to a Phe (Phe 160 calcineurin A). Negligible difference in native PP-1c activity in response to MLR when placing a Phe in position 134. The loss of the hydrogen bond did not have a significant impact on MLR binding.

5. Val 195 PP-1c, Pro 221 calcineurin A. No mutation done, viewed as insignificant.

6. Asp 220 PP-1c, Val 253 calcineurin A. No mutation done, Arg of MLR not important for activity.

7. Val 223 PP-1c mutated to a Cys. No significant difference in PP-1c binding of MLR.

8. PP-1c L7 loop mutated to L7 loop residues of calcineurin A. Simultaneous mutation of Cys 273, Gly 274, Glu 275, Phe 276, and Asp 277 to Leu, Asp, Val, Tyr, and Asn respectively. There was no significant difference from native PP-1c activity in response to MLR concentration.

None of the amino acid differences noted in Table 5 when substituted into PP-1c had any signficant impact on the inhibition of PP-1c by MLR in comparison to native PP-1c. Substitution of the L7 loop from calcineurin A into PP-1c had no impact upon MLR binding to the mutant PP-1c. At this point there seemed to be no explanation to why calcineurin is unaffected by MLR. Further analysis of the comparison between the two crystal structures was needed.

A re-investigation of the comparison between the two crystal structures revealed new evidence which would explain why the mutations of the L7 loop in PP-1c fail to have any significant impact on MLR binding. A previously unnoticed long range structural connection affecting the positioning of the L7 loop in calcineurin A was found.

This connection occurs between the negatively charged Asp 313 and potentially positively charge His 339 in calcineurin A and is outlined in Figure 7. The lack of results in the L7 loop mutation may be rationalized because difference in loop position is externally caused by the interaction in calcineurin A between Asp 313 and His 339 (Figure 7). This connection does not occur in PP-1c because of the lack of the long C-terminal tail resulting in a L7 loop position facilitating MLR binding. The insertion of calcineurin A loop L7 residues would not have changed the L7 loop conformation of PP-1c because the difference in positions is caused by a connection to a residue outside the loop. This difference in L7 loop conformations between PP-1c and calcineurin A may be caused by an ionic and/or hydrogen bonding interaction between His 339 and Asp 313 (closest distances are HE2 of His 339 to OD2 of Asp 313 = 1.98 Å; HE2 of His 339 to OD1 of Asp 313 = 2.57Å) which in calcineurin A causes the L7 loop backbone to be pulled away from the potential MLR binding site relative to the PP-1 position. The result is Leu 312 and Tyr 315 side chains of calcineurin A now occupy the space where MLR should have been able to bind.

Figure 7- *Interaction of His 339 with Asp 313 of the L7 loop in calcineurin A (green)*[5]*. Superimposed onto the backbone C-terminus of calcineurin A is the backbone C-terminus of PP-1c (red) with the attached MLR (blue) in association with the L7 loop of PP-1c. From the prospective of this figure it is clearly seen how the interaction of His 339 with Asp 313 pulls the L7 loop conformation in calcineurin A away from the L7 loop conformation in PP-1c which is required for MLR binding. Thus, the side chains of Leu 312 and Tyr 315 in calcineurin in its L7 loop conformation sterically prevent MLR binding.*

The most plausible hypothesis explaining the difference in L7 loop positions is the interaction between His 339 and Asp 313 (also observed in Figure 5 top left corner where the red section of the bridge of the open space is Asp 313 and the blue section is His 339). This causes the side chains of Leu 312 and Tyr 315 to prevent MLR binding. This hypothesis of calcineurin A resistance to MLR can be further tested by disrupting the association of His 339 and Asp 313 in calcineurin A. This disruption could potentially cause the L7 loop of calcineurin A to adopt a similar conformation to the L7 loop of PP-1c and thereby make this mutant of calcineurin A sensitive to MLR. Such mutagenesis studies of calcineurin A is presently in progress in the Charles Holmes lab.

Figure 8- *Overall view of the superimposition of the backbone PP-1c (red ribbon) and the backbone of calcineurin A (green ribbon). MLR (blue stick) is bound to PP-1c. The L7 loop of PP-1 is highlighted in dark blue while the L7 loop of calcineurin A is highlighted in yellow. The difference in conformations of the L7 loops is hypothesized to be caused by the His 339:Asp 313 interaction highlighted in purple which pulls the calcineurin L7 loop away from the potential MLR binding site resulting in the side chains of Leu 312 and Tyr 315 of calcineurin A sterically preventing MLR binding.*

Another interesting structural feature brought out by the comparison of PP-1c and calcineurin A is why PP-1c is not affected by the immunosuppressants FK506 and cyclosporin A. These immunosuppressants first bind to their respective immunophilins and then interact with both calcineurin A and calcineurin B forming a complex of 4 elements bound together[3]. The reason why PP-1c is unaffected by the immunosuppressants is the lack of the long C-terminal helix that exists in calcineurin A which essentially acts like an anchor for the other elements to bind to (Figure 9).

***Figure 9-** Explanation for why PP-1c is unaffected by immunosuppressants like cyclosporin A and FK 506. PP-1c (backbone is red ribbon) lacks the long, extended C-terminal helix of calcineurin A (backbone is green ribbon) that anchors calcineurin B and the immunosuppressants and immunophilins when forming a complex. MLR bound to PP-1c is in blue and significant side chains that differ in the MLR binding region are displayed as sticks.*

References

1. Wishart, D.S., Boyko, R.F., Willard, L., Richards, F.M., and Sykes, B.D. (1994) SEQSEE: A comprehensive program suite for protein sequence analysis. *Comp. Appl. Biosci* 10:121-132.

2. Goldberg, J., Huang, Hb., Kwon, Yg., Greengard, P., Nairn, A.C., and Kuriyan, J. (1995). Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. *Nature* 376:745-753.

3. Griffith, J.P., Kim, J.L., Kim, E.E., Sintchak, M.D., Thomson, J.A., Fitzgibbon, M.J., Fleming, M.A., Caron, P.R., Hsiao, K., and Navia, M.A. (1995). X-ray structure of calcineurin inhibited by the immunophilin-immunosuppressant FKBP12-FK506 complex. *Cell* 82:507-522.

4. Bagu, J.R., Sykes, B.D., Craig, M.M., and Holmes, C.F.B. (1997). A molecular basis for different interactions of marine toxins with protein phosphatase-1. *J. Biol. Chem.* 272(8):5087-5097.

5. Dawson, J. (1998) Figures in Thesis.

6. Thesis Chapter 6

Structural comparison of the inactive dephospho form and active phospho form of inhibitor-1

Pages 195-197 of this chapter in paper submitted to Journal of Biological Chemistry as McCready, T.L., Craig, M., Bagu, J.R., Sykes, B.D., Semchuk, P., Hodges, R.S., and Holmes, C.F.B. (Submitted).

Other sections of this chapter non-published work by Bagu, J.R.

Summary

Inhibitor-1 is a 18.7 kDa protein that regulates protein phosphatase-1 activity by inhibiting the enzyme when T35 of inhibitor-1 becomes phosphorylated by cAMP-dependent protein kinase. I present a NMR analysis of the dephospho and phospho forms of a completely active N-terminal fragment (residues 1-54) of inhibitor-1. Both the dephospho and phospho forms of the active fragment 1-54 are highly unstructured having a flexible conformation in solution. Phosphorylation of the fragment only results in localized changes centering on T35, and mainly affecting R33, R32, R31, and R30. Comparison of chemical shifts between the 1-54 fragment and a smaller 9-54 fragment of both dephospho and phospho forms indicate that there is an association between I10 and L21 and I29. Using the chemical shift index (CSI) both dephospho and phospho forms of inhibitor-1 show little secondary structure with the only significant secondary structure being from A25-R31. $^{3}J_{NH\alpha CH}$ coupling constants for both dephospho and phospho inhibitor-1 have values between 6-8 Hz for the majority of amino acids indicating that the both fragments are mostly random coil. The addition of the phosphate group to T35 required for inhibition may reflect a similar interaction that microcystin-LR has with R96 and R221 of protein phosphatase-1. Additionally, the strong negative charge of the phosporylated T35 may neutralize in whole or part the positive charges of R30-R33 which may also be useful for binding to protein phosphatase-1.

Introduction

Inhibitor-1 is a 18.7 kDa thermostable protein first described[1] as a specific inhibitor for the catalytic subunit of protein phosphatase-1c (PP-1c)[2]. This inhibitor[3] is one of many endogenous inhibitors controlling PP-1c activity in mammalian cells including DARPP-32[4], inhibitor-2[5], the ribosomal protein RIPP-1[6], the smooth muscle protein kinase C substrate C-kinase activated PP-1 inhibitor of Mr 17,000[7], and the nuclear protein NIPP-1[8] (reviewed in [9]).

Inhibitor-1 is active as a PP-1c inhibitor reducing its dephosphorylation activity only when it is phosphorylated by cAMP-dependent protein kinase at Thr-35[10]. Inhibitor-1 is regulated by reversible phosphorylation being dephosphorylated *in vivo* by protein phosphatase-2A and protein phosphatase-2B (calcineurin)[11-13]. Dephosphorylation of inhibitor-1 to its inactive form by calcineurin is a potential mechanism by which Ca^{2+} can increase PP-1c activity attenuating the effects of cAMP. Activation of PP-1c when inhibitor-1 is inactivated represents a potential phosphatase cascade, with inhibitor-1 lying at a critical junction of two secondary messenger systems[14].

An active fragment (residues 9-54) of inhibitor-1 was isolated[10] and found to retain full inhibitory activity against PP-1c if Thr-35 was phosphorylated. Proteolysis of this fragment into residues 22-54 and 13-41 showed that these fragments were inactive indicating functional importance in residues 9-22 containing many hydrophobic residues. It has been widely proposed that there are two independent structural elements in I-1 that are required for inhibition of PP-1c. These are the phosphorylation site at Thr-35 and residues 9-12 (KIQF)[4,15-17]. Models of I-1 and DARPP-32 bound to the crystal structure of PP-1c have proposed that the four sequential arginine residues preceding Thr-35 interact with acidic amino acids lining a groove located near the active site of PP-1c[16,18-20].

This study encompasses NMR experiments performed on the inactive dephospho inhibitor-1 and active phospho inhibitor-1 for fragment 9-54 and 1-54. Examination of the data reveals that both forms of inhibitor-1 are relative unstructure with phosphorylation of Thr-35 resulting in only a localized change in chemical shifts.

Methods

Fragment 1-54 of inhibitor-1 was made via peptide synthesis. Phosphorylation for the active sample was achieved using a reaction solution of 10ml with 10mM NaGlycerophosphate, 0.4mM EDTA, 0.1mM EGTA, 2mM $MgCl_2$, 9mg ATP, 9mg of dephospho inhibitor-1 fragment 1-54, and 500 units of protein kinase from bovine heart (Sigma Chemical Company P-2645). The reaction solution was left overnight at room temperature. The phosphorylated form of inhibitor-1 was checked by mass spectroscopy and tested in a PP-1c assay and determined to be completely active. The samples used for NMR were 1-2mM inhibitor-1 dissolved in a 10mM potassium phosphate, 50mM sodium chloride buffer with 80-90% H2O/10-20% D2O at pH 6.5. 2,2-dimethyl-2-silapentane-5-sulfonic acid (0.1mM) was added as a ^{1}H NMR chemical shift standard.

^{1}H NMR spectra were recorded at 500MHz and 600MHz on Varian VXR-500 and VXR-600 NMR spectrometers. Proton NMR resonance assignments of both inhibitor-1 peptides dephospho and phospho forms were made using standard sequential assignment methods with double quantum filtered COSY, TOCSY, and NOESY two dimensional ^{1}H NMR spectra. ^{1}H-^{1}H internuclear distance restraints were obtained from two dimensional ^{1}H NOESY taken with a mixing time of 100ms so as to minimize spin diffusion. Assignments were made at 25°C and 5°C. The chemical shift index (CSI) was used to determine any regions of secondary structure based on the alpha-hydrogen chemical shift. Measurements of coupling constants from a two dimensional COSY using J-fit (Author Robert Boyko, University of Alberta) for describing the one-dimensional shapes of the peaks.

Results and Discussion

Chemical Shift Analysis of Inactive Dephospho Inhibitor-1 and Active Phospho Inhibitor-1 Fragment 9-54 and 1-54

The assignments of the hydrogen nuclei in each amino acid residue for the dephospho and phospho forms of the 9-54 fragment has been accomplished (Table 1 and 2 respectively). However, an important residue, I10, was not clearly identified in the 9-54 fragment. Therefore, the fragment 1-54 was synthesized and phosphorylated so that the assignments of the dephospho and phospho forms could be determined with I10 being clearly visible. Table 3 shows the chemical shifts of the inactive dephospho inhibitor-1 fragment 1-54. Table 4 displays the chemical shifts of the active phospho inhibitor-1 fragment 1-54. The most difficult region to assign was residues R30-R33, encompassing 4 continuous arginines.

Table 1-Chemical shifts of inhibitor-1 dephospho residues 9-54 at 5°C (Values in ppm)

Residue	HN	α	β	β'	γ	γ'	δ	δ'	ε	ε'
K9	N/A	4.03	1.88		1.44	1.38	1.70		2.99	
I10	8.70	4.12	1.78		1.45	1.15	0.79			
Q11	8.63	4.34	1.89		2.25				7.59	6.97
F12	8.46	4.70	3.13	3.02			7.24		7.33	
T13	8.23	4.31	4.05		1.14					
V14	8.34	4.35	2.09		0.99					
P15		4.37	2.29	1.83	2.06		3.91	3.71		
L16	8.44	4.30	1.61	1.57	1.53		0.91	0.87		
L17	8.37	4.38	1.60		1.54		0.92	0.86		
E18	8.46	4.57	2.01	1.87	2.26					
P19		4.36	2.25	1.84	2.01		3.78	3.72		
H20	8.70	4.63	3.14					7.14		
L21	8.36	4.35	1.57	1.53	1.48		0.82			
D22	8.57	4.85	2.82	2.65						
P23		4.34	2.35	1.96	2.07		3.91	3.89		
E24	8.50	4.22	2.09	2.00	2.31	2.27				
A25	7.96	4.20	1.48							
A26	8.37	4.12	1.44							
E27	8.25	4.19	2.08	2.04	2.34	2.28				
Q28	8.08	4.10	2.11	2.06	2.47	2.41			7.70	6.98
I29	7.91	3.86	1.89		1.59	1.18	0.90			

Residue	HN	α	β	β'	γ	γ'	δ	δ'	ε	ε'
R30	8.10	4.15	1.88		1.71	1.60	3.22		7.50	
R31	8.12	4.23	1.91	1.83	1.76	1.66	3.19		7.47	
R32	8.04	4.29	1.88	1.82	1.75	1.66	3.19		7.31	
R33	8.31	4.61	1.87	1.79	1.72		3.22		7.30	
P34		4.50	2.30	1.88	2.03		3.85	3.64		
T35	8.55	4.57	4.15		1.28					
P36		4.37	2.33	1.98	2.06		3.90	3.72		
A37	8.58	4.31	1.41							
T38	8.21	4.26	4.16		1.21					
L39	8.36	4.37	1.62		1.58		0.87			
V40	8.32	4.06	2.02		0.93					
L41	8.60	4.49	1.70	1.64	1.59		0.94	0.86		
T42	8.33	4.39	4.29		1.22					
S43	8.49	4.46	3.93	3.83						
D44	8.49	4.61	2.72	2.68						
Q45	8.40	4.36	2.00	1.95	2.38	2.19			7.64	6.95
S46	8.47	4.45	3.88							
S47	8.50	4.79	3.90	3.83						
P48		4.44	2.29	1.91	2.02		3.80	3.74		
E49	8.64	4.26	2.02	1.92	2.32	2.26				
V50	8.32	4.11	2.08		0.93					
D51	8.54	4.61	2.71	2.58						
E52	8.52	4.27	2.03	1.93	2.26	2.22				
D53	8.56	4.60	2.75	2.60						
R54	7.87	4.16	1.84	1.75	1.70	1.58	3.20		7.29	

Table 2-Chemical shifts of inhibitor-1 phospho residues 9-54 at 5°C (Values in ppm)

Residue	HN	α	β	β'	γ	γ'	δ	δ'	ε	ε'
K9	N/A	4.02	1.88		1.44	1.38	1.68		2.98	
I10	8.70	4.12	1.78		1.46	1.14	0.79			
Q11	8.63	4.32	1.89		2.25				7.60	6.96
F12	8.49	4.70	3.13	3.02			7.24		7.33	
T13	8.23	4.31	4.05		1.14					
V14	8.33	4.35	2.09		0.99					
P15		4.37	2.27	1.84	2.06		3.90	3.71		
L16	8.43	4.30	1.61	1.57	1.53		0.91	0.88		
L17	8.36	4.39	1.60		1.54		0.92	0.85		
E18	8.45	4.57	2.01	1.87	2.26					
P19		4.36	2.24	1.84	2.01		3.78	3.71		
H20	8.69	4.63	3.14					7.14	8.24	
L21	8.36	4.34	1.57	1.53	1.48		0.83			
D22	8.57	4.86	2.81	2.61						
P23		4.37	2.34	1.97	2.09		3.91	3.89		
E24	8.51	4.21	2.09	2.00	2.32	2.26				
A25	7.99	4.21	1.46							
A26	8.34	4.16	1.42							
E27	8.27	4.18	2.06	2.02	2.32	2.28				
Q28	8.15	4.16	2.12	2.06	2.44	2.40			7.68	6.98
I29	8.00	3.97	1.88		1.55	1.20	0.90			
R30	8.22	4.23	1.85		1.68	1.61	3.19		7.44	
R31	8.22	4.24	1.84		1.69	1.63	3.18		7.43	
R32	8.18	4.27	1.81		1.76	1.63	3.19		7.39	
R33	8.57	4.61	1.86	1.79	1.72		3.22		7.40	
P34		4.40	2.30	1.86	2.06		3.87	3.64		
PhosphoT35	9.51	4.38	4.31		1.35					
P36		4.37	2.33	1.94	2.08		4.05	3.73		
A37	8.64	4.32	1.41							
T38	8.27	4.25	4.14		1.20					
L39	8.39	4.38	1.63		1.57		0.87			
V40	8.34	4.06	2.02		0.93					
L41	8.60	4.49	1.70	1.64	1.59		0.94	0.86		
T42	8.33	4.39	4.29		1.22					

Residue	HN	α	β	β'	γ	γ'	δ	δ'	ε	ε'
S43	8.49	4.47	3.93	3.84						
D44	8.50	4.61	2.72	2.68						
Q45	8.40	4.37	1.99	1.95	2.37	2.18			7.63	6.94
S46	8.46	4.45	3.88							
S47	8.50	4.80	3.89	3.83						
P48		4.44	2.30	1.91	2.03		3.80	3.74		
E49	8.64	4.26	2.03	1.92	2.32	2.26				
V50	8.32	4.12	2.08		0.93					
D51	8.54	4.61	2.72	2.58						
E52	8.47	4.19	2.03	1.93	2.32	2.23				
D53	8.56	4.64	2.76	2.60						
R54	7.86	4.16	1.84	1.75	1.70	1.58	3.20		7.32	

Table 3-Chemical shifts of inhibitor-1 dephospho residues 1-54 at 5°C (Values in ppm)

Residue	HN	α	β	β'	γ	γ'	δ	δ'	ε	ε'
M1	8.51	4.44	2.07	1.98	2.61	2.56			1.43	
E2	8.75	4.28	2.06	1.96	2.31	2.27				
Q3	8.58	4.30	2.10	1.98	2.37	2.33			7.70	7.23
D4	8.52	4.58	2.76	2.66						
N5	8.58	4.78	2.88	2.77			7.69	6.98		
S6	8.31	4.67	3.89							
P7		4.43	2.29	1.88	2.03		3.84	3.70		
R8	8.50	4.28	1.80	1.74	1.64		3.19		7.25	
K9	8.48	4.29	1.78		1.44	1.38	1.68		2.98	
I10	8.36	4.08	1.78		1.45	1.15	0.79			
Q11	8.54	4.32	1.89		2.25				7.61	6.97
F12	8.46	4.70	3.13	3.02			7.24		7.33	
T13	8.23	4.31	4.05		1.14					
V14	8.34	4.35	2.09		0.99					
P15		4.37	2.29	1.83	2.06		3.91	3.71		
L16	8.44	4.30	1.61	1.57	1.53		0.91	0.87		
L17	8.37	4.38	1.60		1.54		0.92	0.86		
E18	8.46	4.57	2.01	1.87	2.26					
P19		4.36	2.25	1.84	2.01		3.78	3.72		
H20	8.63	4.63	3.14					7.14		

Residue	HN	α	β	β'	γ	γ'	δ	δ'	ε	ε'
L21	8.27	4.35	1.57	1.53	1.48		0.82			
D22	8.53	4.85	2.82	2.65						
P23		4.34	2.35	1.96	2.07		3.91	3.89		
E24	8.50	4.22	2.09	2.00	2.35	2.27				
A25	7.96	4.20	1.48							
A26	8.37	4.12	1.44							
E27	8.25	4.19	2.08	2.04	2.34	2.28				
Q28	8.04	4.10	2.15	2.10	2.47	2.41			7.70	6.98
I29	7.87	3.86	1.89		1.59	1.18	0.90			
R30	8.06	4.15	1.88		1.71	1.60	3.22		7.55	
R31	8.08	4.23	1.91	1.83	1.76	1.66	3.19		7.51	
R32	7.98	4.29	1.88	1.82	1.75	1.66	3.19		7.31	
R33	8.31	4.61	1.87	1.79	1.72		3.22		7.30	
P34		4.50	2.30	1.88	2.03		3.85	3.64		
T35	8.55	4.57	4.15		1.28					
P36		4.37	2.33	1.98	2.06		3.90	3.72		
A37	8.58	4.31	1.41							
T38	8.21	4.26	4.16		1.21					
L39	8.36	4.37	1.62		1.58		0.87			
V40	8.32	4.06	2.02		0.93					
L41	8.60	4.49	1.70	1.64	1.59		0.94	0.86		
T42	8.33	4.39	4.29		1.22					
S43	8.49	4.46	3.93	3.83						
D44	8.49	4.61	2.72	2.68						
Q45	8.40	4.36	2.00	1.95	2.38	2.19			7.64	6.95
S46	8.47	4.45	3.88							
S47	8.50	4.79	3.90	3.83						
P48		4.44	2.29	1.91	2.02		3.80	3.74		
E49	8.64	4.26	2.02	1.92	2.32	2.26				
V50	8.32	4.11	2.08		0.93					
D51	8.54	4.61	2.71	2.58						
E52	8.52	4.19	2.03	1.93	2.26	2.22				
D53	8.56	4.60	2.75	2.64						
R54	8.40	4.28	1.96	1.75	1.70	1.66	3.20		7.33	

Table 4-Chemical shifts of inhibitor-1 phospho residues 1-54 at 5°C (Values in ppm)

Residue	HN	α	β	β'	γ	γ'	δ	δ'	ε	ε'
M1	8.50	4.44	2.08	1.98	2.61	2.56				
E2	8.74	4.27	2.05	1.96	2.31	2.27				
Q3	8.57	4.29	2.10	1.98	2.37	2.33			7.69	7.22
D4	8.51	4.58	2.75	2.65						
N5	8.56	4.79	2.87	2.76	7.68	6.98				
S6	8.30	4.66	3.89							
P7		4.42	2.30	1.89	2.03		3.84	3.69		
R8	8.49	4.29	1.81	1.74	1.64		3.20		7.25	
K9	8.47	4.30	1.78		1.44	1.38	1.68		2.98	
I10	8.35	4.08	1.78		1.46	1.14	0.79			
Q11	8.53	4.32	1.89		2.25				7.60	6.96
F12	8.45	4.70	3.13	3.02			7.24		7.33	
T13	8.23	4.31	4.05		1.14					
V14	8.33	4.35	2.09		0.99					
P15		4.37	2.27	1.84	2.06		3.90	3.71		
L16	8.43	4.30	1.61	1.57	1.53		0.91	0.88		
L17	8.36	4.39	1.60		1.54		0.92	0.85		
E18	8.45	4.57	2.01	1.87	2.26					
P19		4.36	2.24	1.84	2.01		3.78	3.71		
H20	8.62	4.63	3.14					7.14	8.24	
L21	8.28	4.34	1.57	1.53	1.48		0.83			
D22	8.53	4.86	2.81	2.61						
P23		4.37	2.34	1.97	2.09		3.91	3.89		
E24	8.51	4.21	2.09	2.00	2.32	2.26				
A25	7.99	4.21	1.46							
A26	8.34	4.16	1.42							
E27	8.27	4.18	2.06	2.02	2.32	2.28				
Q28	8.15	4.16	2.12	2.06	2.44	2.40			7.68	6.98
I29	7.96	3.97	1.88		1.55	1.20	0.90			
R30	8.17	4.23	1.85		1.68	1.61	3.19		7.44	
R31	8.17	4.24	1.84		1.69	1.63	3.18		7.43	
R32	8.18	4.27	1.81		1.76	1.63	3.19		7.31	
R33	8.57	4.61	1.86	1.79	1.72		3.22		7.40	

Residue	HN	α	β	β'	γ	γ'	δ	δ'	ε	ε'
P34		4.40	2.30	1.86	2.06		3.87	3.64		
PhosphoT35	9.63	4.38	4.31		1.35					
P36		4.37	2.33	1.94	2.08		4.05	3.73		
A37	8.64	4.32	1.41							
T38	8.27	4.25	4.14		1.20					
L39	8.39	4.38	1.63		1.57		0.87			
V40	8.34	4.06	2.02		0.93					
L41	8.60	4.49	1.70	1.64	1.59		0.94	0.86		
T42	8.33	4.39	4.29		1.22					
S43	8.49	4.47	3.93	3.84						
D44	8.50	4.61	2.72	2.68						
Q45	8.40	4.37	1.99	1.95	2.37	2.18			7.63	6.94
S46	8.46	4.45	3.88							
S47	8.50	4.80	3.89	3.83						
P48		4.44	2.30	1.91	2.03		3.80	3.74		
E49	8.64	4.26	2.03	1.92	2.32	2.26				
V50	8.32	4.12	2.08		0.93					
D51	8.54	4.61	2.72	2.58						
E52	8.52	4.19	2.03	1.93	2.27	2.23				
D53	8.56	4.60	2.76	2.64						
R54	8.40	4.28	1.96	1.75	1.70	1.66	3.20		7.32	

Comparison of 9-54 fragments with 1-54 fragments of inhibitor-1

Chemical shifts comparsons can be made between the 9-54 and 1-54 fragment and between the dephospho and phospho forms of each fragment. There are only minor changes (greater than 0.03ppm) when comparing the chemical shifts of the 9-54 fragment with residues 9-54 in the 1-54 fragment for both dephospho (Table 5) and phospho forms (Table 6). The first set of changes in the chemical shifts are residues K9-Q11 caused by the elongation of the amino acid chain at that N-terminus. Extention of the N-terminus also caused the chemical shifts to change in residues H20-D22 and Q28-R32. This indicates that residues K9-Q11 are in proximity to H20-D22 and Q28-R32. This interaction probably is hydrophobic in nature occurring between the essential residue I10 with L21 and I29. The final set of changes in the C-terminus reflects different procedures used to synthesize the peptides. The 9-54 fragment C-terminus was left with a negative charge in the carboxyl group of R54, while the 1-54 fragment had an uncharged amide group attached to the C-terminal R54. These conclusions are true for both forms of the peptide.

Table 5-Comparison of fragment 9-54 with fragment 1-54 chemical shifts of inhibitor-1 dephospho form (Values in ppm, only differences greater than 0.03 ppm are shown)

Atom	9-54 Fragment	1-54 Fragment	Difference
K9 HN	N/A	8.48	N/A
K9 HA	4.03	4.29	+0.26
K9 HB	1.88	1.78	-0.10
I10 HN	8.70	8.36	-0.34
I10 HA	4.12	4.08	-0.04
Q11 HN	8.63	8.54	-0.09
H20 HN	8.70	8.63	-0.07
L21 HN	8.36	8.27	-0.09
D22 HN	8.57	8.53	-0.04
E24 HG1	2.31	2.35	+0.04
Q28 HN	8.08	8.04	-0.04
Q28 HB1	2.11	2.15	+0.04
Q28 HB2	2.06	2.10	+0.04
I29 HN	7.91	7.87	-0.04
R30 HN	8.10	8.06	-0.04
R30 HE	7.50	7.55	+0.05
R31 HN	8.12	8.08	-0.04
R31 HE	7.47	7.51	+0.04
R32 HN	8.04	7.98	-0.06
E52 HA	4.27	4.19	-0.08
D53 HB2	2.60	2.64	+0.04
R54 HN	7.87	8.40	+0.53
R54 HA	4.16	4.28	+0.12
R54 HB1	1.84	1.96	+0.12
R54 HG2	1.58	1.66	+0.08
R54 HE	7.29	7.33	+0.04

Table 6-Comparison of fragment 9-54 with fragment 1-54 chemical shifts of inhibitor-1 phospho form (Values in ppm, only differences greater than 0.03 ppm are shown)

Atom	9-54 Fragment	1-54 Fragment	Difference
K9 HN	N/A	8.47	N/A
K9 HA	4.02	4.30	+0.28
K9 HB	1.88	1.78	-0.10
I10 HN	8.70	8.35	-0.35
I10 HA	4.12	4.08	-0.04
Q11 HN	8.63	8.53	-0.10
F12 HN	8.49	8.45	-0.04
H20 HN	8.69	8.62	-0.07
L21 HN	8.36	8.28	-0.08
D22 HN	8.57	8.53	-0.04
I29 HN	8.00	7.96	-0.04
R30 HN	8.22	8.17	-0.05
R31 HN	8.22	8.17	-0.05
R32 HE	7.39	7.31	-0.08
T35 HN	9.51	9.63	+0.12
E52 HA	4.27	4.19	-0.08
E52 HG1	2.32	2.27	-0.05
D53 HA	4.64	4.60	-0.04
D53 HB2	2.60	2.64	+0.04
R54 HN	7.86	8.40	+0.54
R54 HA	4.16	4.28	+0.12
R54 HB1	1.84	1.96	+0.12
R54 HG2	1.58	1.66	+0.08

Analysis of phosphorylation of inhibitor-1 fragment 1-54

A comparison of chemical shifts between the dephospho form and phospho forms of inhibitor-1 1-54 provides an explanation of why the dephospho form is inactive while the phospho form is an active inhibitor of PP-1c. The chemical shifts are almost identical in both forms except for some proton nuclei in residues 28-38 (Table 7). Therefore, the phosphorylation of inhibitor-1 causes only a localized change in the fragment in the area of the phosphorylation site T35. Note that the chemical shift change is not centered around T35 but is more pronounced in the residues before T35. This could indicated a minor structural change. However, in the case of R30 and R33 (show the greatest changes besides T35), the neutralization of their positive charges by the addition of the strong negatively charged phosphate group could cause the difference in chemical shifts between the two forms. There are no significant changes in the NOE crosspeaks between the two forms indicating that there is no major structural changes induced by phosphorylation. The neutralization of the positively charged arginines (particularly R33 and R32) may be crucial for binding to PP-1c.

Table 7-Chemical shift changes due to phosphorylation of inhibitor-1 fragment 1-54 (Values in ppm, only differences greater than 0.03 ppm are shown)

Atom	Dephospho	Phospho	Difference
D22 HB2	2.65	2.61	-0.04
Q28 HN	8.04	8.15	+0.11
Q28 HA	4.10	4.16	+0.06
Q28 HB2	2.10	2.06	-0.04
I29 HN	7.87	7.96	+0.09
I29 HA	3.86	3.97	+0.11
I29 HG1	1.59	1.55	-0.04
R30 HN	8.06	8.17	+0.11
R30 HA	4.15	4.23	+0.08
R30 HE	7.55	7.44	-0.11
R31 HN	8.08	8.17	+0.09
R31 HG1	1.76	1.69	-0.07
R31 HE	7.51	7.43	-0.08
R32 HN	7.98	8.18	+0.20
R33 HN	8.31	8.57	+0.26
R33 HE	7.30	7.40	+0.10

Atom	Dephospho	Phospho	Difference
P34 HA	4.50	4.40	-0.10
T35 HN	8.55	9.63	+1.08
T35 HA	4.57	4.38	-0.19
T35 HB	4.15	4.31	+0.16
T35 HG	1.28	1.35	+0.07
P36 HB2	1.98	1.94	-0.04
P36 HD1	3.90	4.05	+0.15
A37 HN	8.58	8.64	+0.06
T38 HN	8.21	8.27	+0.06

Using the Chemical Shift Index (CSI) to analyze the chemical shifts of dephospho and phospho inhibitor-1 fragment 1-54

The chemical shift index (CSI) gives insights into the secondary structure of proteins based on their alpha proton chemical shifts. Using this technique the chemical shift index results are displayed in Table 8 for both the dephospho and phospho forms. The same results occur in the residues that are duplicated in the 9-54 fragment. Other than the CSI index no other information shows beta strands from P15 to E18 and L39 to L41. There may be a helix, or a loose series of loops between residues A25-R32 which may also encompass D22 which shows strong NOE connections to A25. Other than this helix or series of loop there is no significant secondary structure.

Table 8-Chemical shift Index analysis of HA protons for dephospho and phospho inhibitor-1 fragment 1-54

Dephospho		Phospho	
Secondary Struct.	Residues	Secondary Struct.	Residues
Coil	M1-V14	Coil	M1-V14
Beta Strand	P15-E18	Beta Strand	P15-E18
Coil	P19-E24	Coil	P19-E24
Helix	A25-R31	Helix	A25-R32
Coil	R32-T38	Coil	R33-T38
Beta Strand	L39-L41	Beta Strand	L39-L41
Coil	T42-R54	Coil	T42-R54

Coupling constant analysis from 2D-dqf Cosy of dephospho and phospho inhibitor-1 fragment 1-54

The final data which describes both dephospho and phospho forms of inhibitor-1 as being relatively unstructured and flexible is the analysis of the coupling constants. Coupling constants from double quantum filtered COSYs were measured so as to view phi angle restraints (HN-N-CA-HA). Low coupling constant values (below 6) usually indicate a helical secondary structure motif. High coupling constant values (above 8) normally imply beta-strand secondary structure. Intermediate values (6-8) are indicative of random coil secondary structure. Table 9 is the results of measuring the coupling constants from a two dimensional COSY using J-fit (Author Robert Boyko, University of Alberta) from the ω1 trace. Residues not mentioned in Table 9 are excluded because spectral overlap did not permit accurate measurements of the coupling constant (e.g. R30-R33) or in the case of proline there is no possible coupling constant to measure. As observed in Table 7 most coupling constants fall in between 6-8 indicating the dephospho and phospho inhibitor-1 fragments have little secondary structure. This combined with the lack of long range NOE's suggest there is little secondary or tertiary structure, and inhibitor-1 must be flexible in solution. Also note that no improvement to the spectra in terms of significant chemical shift changes or additional long range NOE's occurred when adjusting pH, temperature, and solvents (including TFE). These fragments simply refused to show any structure!

Table 9-Coupling constant analysis of inhibitor-1 dephospho and phospho forms fragment 1-54 (Values in Hz)

Residue	Dephospho	Phospho
M1	6.5	6.7
E2	6.2	6.2
Q3	6.5	6.7
D4	6.5	7.4
N5	7.9	8.3
S6	5.8	6.1
I10	7.6	Overlapped
Q11	6.8	6.9
F12	7.9	7.2

Residue	Dephospho	Phospho
T13	7.3	7.3
V14	7.5	8.3
L16	6.6	6.0
L17	Overlapped	6.7
E18	7.0	6.7
H20	7.9	7.7
L21	7.3	7.5
E24	Overlapped	5.9
A25	4.6	6.4
E27	4.1	Overlapped
Q28	7.1	7.4
I29	8.3	5.4
T35	6.7	6.2
A37	6.5	6.0
T38	6.7	6.5
L39	Overlapped	8.6
V40	7.5	9.0
L41	6.1	6.1
T42	7.5	7.5
S43	6.0	6.8
D44	6.6	6.5
Q45	7.2	9.0
S46	6.2	6.5
S47	7.3	7.3
E49	6.4	6.2
V50	7.1	7.7
D51	Overlapped	6.4
E52	5.4	5.4
D53	Overlapped	6.2
R54	7.3	7.4

Comparison to previously studied dephospho verses phospho proteins

Generally, there are 2 possible effects that phosphorylation has on a protein. The first possibility is that phosphorylation results in no major structural change, however the electrostatic nature of the protein is dramatically effected or the addition of the phosphate group. An example of the first possibility is isocitrate dehydrogenase, an enzyme involved in the citric acid cycle. When energy is abundant isocitrate is split by isocitrate dehydrogenase in a dephosphorylated form into succinate and glyoxylate, a key reaction of the citric acid cycle. When energy is required, the activity of isocitrate dehydrogenase is turned off by phosphorylation at S113 which blocks the binding of isocitrate to the enzyme[21,22]. The consequence of this is that isocitrate is now oxidatively decarboxylated to α-ketoglutarate which is part of the glyoxylate cycle. Thus, the competition for isocitrate by the citric acid cycle and the glyoxylate cycle is mediated by the phosphorylation state of isocitrate dehydrogenase which in turn is regulated by whether energy is abundant or required. Another example is the histidine-containing protein, an intermediate protein in the bacterial phosphoenolpyruvate:sugar phosphotransferase system. Phosphorylation of this protein results in no major structural changes, but does allow for favorable electrostatic interactions allowing for further biological activity[23]. Inhibitor-1 appears to be like the histidine-containing protein and isocitrate dehydrogenase in that the addition of the phosphate group by phosphorylation has no major structural consequences but instead has a minor structural impact by the addition of the phosphate group and major electrostatic changes favorable to interacting with PP-1c (e.g. addition of strong negative charge and neutralization of R30-33 positive charges).

The second possibility is that phosphorylation results in a change in structural conformation of the protein. The change in structural conformation could be at the phosphorylation site or allosteric affecting another distant part of the protein. The best example of this is glycogen phosphorylase which catalyzes the first step in glycogen degradation to generate energy. Phosphorylation controls the reversible step converting dephosphorylated phosphorylase b (inactive) to phosphorylated phosphorylase a (active). This is achieved by phosphorylation at S14 of the phosphorylase dimer. The consequence is a major structural change at the subunit interface having long-range effects causing glycogen phosphorylase to become almost fully active by phosphorylation[24-27].

Implications for bound form of inhibitor-1

The bound form of inhibitor-1 is required to have the strong negative charge of the phosphate group attached to T35 in order for inhibition of PP-1c to take place. It might be possible that T35 when phosphorylated could interact with R96 and R221 of PP-1c in a similar manner to the Masp and Glu carboxyl groups of microcystin-LR, thus preventing PP-1c activity by interacting with some residues of the catalytic site of PP-1c. However, it would also be of interest to observe how the 4 arginines of inhibitor-1 are affected by the addition of the phosphate group and whether this interupts any potential interaction with PP-1c residues.

References

1. Huang, F.L. and Glinsmann, W.H. (1976) Eur J Biochem 70, 419-26.

2. Cohen, P. (1978) Curr. Top. Cell. Reg. 14, 117-196.

3. Nimmo, G.A. and Cohen, P. (1978) Eur J Biochem 87, 341-51.

4. Hemmings, H.C., Nairn, A.C., Elliott, J.I. and Greegard, P. (1990) J Biol Chem 265, 20369-76.

5. Park, I.K. and DePaoli-Roach, A.A. (1994) J Biol Chem 269, 28919-28.

6. Beullens, M., Stalmans, W. and Bollen, M. (1996) Eur J Biochem 239, 183-189.

7. Eto, M., Ohmori, T., Suzuki, M., Furuya, K. and Morita, F. (1995) J Biochem 118, 1104-7.

8. Van Eynde, A., Wera, S., Beullens, M., Torrekens, S., Van Leuven, F., Stalmans, W. and Bollen, M. (1995) J Biol Chem 270, 28068-74.

9. Connor, J.H. et al. (1998) J Biol Chem 273, 27716-24.

10. Aitken, A., Bilham, T. and Cohen, P. (1982) Eur J Biochem 126, 235-46.

11. Cohen, P. (1989) Annu Rev Biochem 58, 453-508.

12. Hubbard, M.J. and Cohen, P. (1989) Eur J Biochem 186, 711-720.

13. Shenolikar, S. and Nairn, A.C. (1991) Adv Secon Messenger Phosphoprotein Res 23. 1-121.

14. Cohen, P. and Cohen, P.T.W. (1989) J Biol Chem 36, 21435-21438.

15. Endo, S., Zhou, X., Connor, J., Wang, B. and Shenolikar, S. (1996) Biochemistry 35, 5220-8.

16. Kwon, Y.G., Huang, H.B., Desdouits, F., Girault, J.A., Greengard, P. and Nairn, A.C. (1997) Proc Natl Acad Sci USA 94, 3536-41.

17. Desdouits, F. et al. (1995) Biochem Biophys Res Commun 206, 653-658.

18. Connor, J.H., Quan, H., Oliver, C. and Shenolikar, S. (1998) Methods Mol Biol 93, 41-58.

19. Goldberg, J., Huang, H.B., Kwon, Y.G., Greengard, P., Nairn, A.C. and Kuriyan, J. (1995) Nature 376, 745-53.

20. Barford, D., Das, A.K. and Egloff, M.P. (1998) Annual Reviews in Biophysical and Biomolecular Structure 27, 133-64.

21. Hurley, J.H., Thorsness, P.E., Ramalingam, V., Helmers, N.H., Koshland, D.E., and Stroud, R.M. (1989) Proc. Natl. Acad. Sci, USA 86, 8635-8639.

22. Hurley, J.H., Dean, A.M., Thorsness, P.E., Koshland, D.E., and Stroud, R.M. (1990) J Biol Chem 265, 3599-3602.

23. Rajagopal, P., Waygood, E.B., Klevit, R.E. (1994) Biochemistry 33, 15271-15282.

24. Sprang, S.R., Acharya, K.R., Goldsmith, E.J., Stuart, D.I., Varvill, K., Fletterick, R.J., Madsen, N.B., and Johnson, L.N. (1988) Nature 336, 215-221.

25. Sprang, S.R., Withers, S.G., Goldsmith, E.J., Fletterick, R.J., and Madsen, N.B. (1991) Science 254, 1367-1371.

26. Barford, D., Hu, S.H., and Johnson, L.N. (1991) J Mol Biol 218, 233-260.

27. Rath, V.L. and Fletterick, R.J. (1994) Nature Structural Biology 1, 681-690.

Thesis Chapter 7

$^{15}N/^{13}C$ Labeled microcystin-LR bound to protein phosphatase-1c

This chapter contains non-published work. NMR spectra by Bagu, J.R.. PP-1c expression and purification by Luu, H.A. and Bagu, J.R.. Labeled microcystin-LR production by Dr. Klaus Frobel, Dr. Hartwig Muller, and Dr. Thomas Henkel.

Summary

In order to determine the bound structure of microcystin-LR when complexed to protein phosphatase-1c ^{15}N and ^{13}C experiments were performed on a double labelled sample of microcystin-LR. The ^{15}N and ^{13}C chemical shifts were determined for microcystin-LR in the free state. However, the ^{15}N and ^{13}C chemical shifts for bound microcystin-LR could not be determined due to solubility difficulties with protein phosphatase-1c. Further investigation into increasing the solubility of protein phosphatase-1c is required. Additionally, experiments on a higher field spectrometer (greater than 600 MHz) may result in acceptable spectra being acquired.

Introduction

The microcystin class of peptide hepatotoxins are metabolites of cyanobacteria in the genera *Microcystis* and *Anabaena* that grow worldwide in fresh and brackish water[1]. These toxins are responsible for extensive wildlife fatalities, and adverse effects on human health have also been recognized in countries where drinking water supplies contain cyanobacteria. It has been shown that microcystin-LR is a potent inhibitor of the catalytic subunits of protein phosphatase-1 and -2A (PP-1c and PP-2Ac) as well as a powerful tumor promoter[2-6]. Inhibition of these enzymes in the liver is apparently associated with hepatocyte deformation due to reorganization of microfilaments[7]. Liver tumor promotion may be linked to the ability of this cyclic peptide to promote hyperphosphorylation of cytokeratins associated with morphological changes in rat hepatocytes[8].

Previously, the average free solution structure of microcystin-LR[9] was compared to the bound crystal structure[10] of microcystin-LR:PP-1c (Chapter 3). Comparison of these structures revealed a similar three-dimensional conformation. Double $^{15}N/^{13}C$ labelled microcystin-LR (acquired from Bayer in Germany: Dr. Klaus Frobel, Dr. Hartwig Muller, and Dr. Thomas Henkel; GB Pharma Research, Chemical Reasearch NASP) bound to PP-1c and was examined by NMR in an attempt to confirm the bound conformation of microcystin-LR in solution. ^{15}N- and ^{13}C- HSQC spectra were acquired for both free and bound microcystin-LR. ^{15}N and ^{13}C assignments were made for the free form of microcystin-LR. Poor spectra quality of the bound form made it impossible to assign peaks.

Methods

Recombinant PP-1c expression and purification

Hue Anh Luu is thanked for helping me learn this expression and purification protocol and for working me on the very successful preparations. Underlined sentences represent areas where I contributed to the protocol which resulted in production from typically to 10 mg to 100 mg of PP-1c. The preparation of PP-1c can be divided into different sections, with each section typically representing a 1 day of work.

Part 1

A. Transform *E. coli* DH5α cells with wild-type cDNA in a pCW expression vector. Inoculate 400ml of Luria-Bertani medium containing 100*u*g/ml ampicillin (use 200ul stock solution) and 1mM $MnCl_2$ (use 2ml stock solution). Grow overnight at 37°C.

B. Prepare stock solution of 10ml 200mg/ml ampicillin that has to be filtered with autoclaved water. 200mg/ml * g/1000ml * 10ml = 2g.

C. Prepare stock solution of isopropyl *B*-D-thiogalactopyranoside (IPTG) that has to be filtered with autoclaved water. Use a 1g bottle. 1g * mol/238.3g * L/0.5mol = 0.0084L or 8.4ml.

D. Prepare stock solution of 10ml 1*u*l/ml 0.1% vitamin B1 that has to be filtered with autoclaved water. Need 4ml (4 X 1ml) for a 4L (4 X 1L) medium. 1% is 0.1g/100ml * 10ml = 0.01g.

E. Prepare stock solution of 200ml 200mM $MnCl_2$ that has to be autoclaved. 0.2mol/L * 197.9g/mol * 0.2L = 7.916g.

F. Prepare Luria-Bertani media (5 X 1L) by adding 25g to 1L in a flask. Autoclave with the $MnCl_2$. Use 4 for second round of growth with the other for initial growth.

Part 2

A. After overnight growth, use culture to inoculate 4L (4 X 1L) Luria-Bertani medium containing 1mM $MnCl_2$, 100*u*g/ml ampicillin, 1*u*l/ml of 0.1% vitamin B1 and grow at 37°C until absorbance at 600nm is at least 0.3. (Note: first time 33°C for 1 hour gave absorbances of 0.457, 0.497, 0.477, and 0.517). Have stock $MnCl_2$, want 0.001mol/L *

197.9g/mol * 4L = 0.7916g, so use 0.7916g * mol/197.9g * L/0.2mol = 0.02L or 20ml stock. For stock ampicillin, want 100*u*g/ml * 1000ml/L * 4L = 400mg, so use 400mg * ml/200mg = 2ml stock. For stock vitamin B1, want 0.1% of 4L so use 4ml of stock. Induce expression from the plasmid overnight at 30°C by addition of 1mM IPTG. Have stock IPTG, want 0.001mol/L * 238.3g/mol * 4L = 0.9532g, so use 0.9532g * mol/238.3g * L/0.5mol = 0.008L or 8ml of stock.

B. Prepare 1L stock solution of 10X buffer A without DTT, $MnCl_2$, PMSF, benzamidine, and glycerol. Buffer A is 50mM imidazole, 100mM NaCl, 0.5mM EDTA, 0.5mM EGTA, 2mM $MnCl_2$, 3mM DTT, 0.5mM PMSF, 2mM benzamidine, and 10% glycerol at pH 7.5. Therefore, want 500mM imidazole, 1M NaCl, 5mM EDTA, and 5mM EGTA. 0.5mol/L * 68.08g/mol * 1L = 34.04g imidazole. 1mol/L * 58.44g/mol * 1L = 58.44g NaCl. 0.005mol/L * 372.24g/mol * 1L = 1.8612g EDTA. 0.005mol/L * 380.4g/mol * 1L = 1.902g EGTA.

C. Autoclave centrifuge containers.

Part 3

A. Spin cultures approximately 1L at a time at 4000rpm for approximately 30 minutes. Keep pellet and discard supernatant. Can store pellet at -70°C.

B. Prepare 4L of 1X buffer A by diluting 10X stock buffer A and adding 2mM $MnCl_2$, 3mM DTT, and 10% glycerol (but no PMSF or benzamidine). This means that 400ml 10X stock buffer A is used. 400ml of glycerol is used. For $MnCl_2$, want 0.002mol/L * 197.9g/mol * 4L = 1.5832g, so use 1.5832g * mol/197.9g * L/0.2mol = 0.04L or 40ml of stock solution. For DTT need 0.003mol/L * 154.24g/mol * 4L = 1.8509g. When PMSF is to be added, want 0.0005mol/L * 174.19g/mol * 2L = 0.1742g, so use 0.1742g * mol/174.19g * L/0.75mol = 0.0013L or 1.3ml of stock solution (add 0.667ml for separate 1L buffer with salt and extraction buffer). When benzamidine is to be added, want 0.002mol/L * 156.6g/mol * 2L = 0.6264g, so use 0.6264g * mol/156.6g * L/1mol = 0.004L or 4ml of stock solution (add 2ml for separate 1L buffer with salt and extraction buffer).

C. Prepare stock solution 0.75M phenylmethylsufonyl fluoride (PMSF) 20ml in methanol. 0.75mol/L * 174.19g/mol * 0.02L = 2.6129g.

D. Prepare stock solution 1M benzamidine 20ml. 1mol/L * 156.6g/mol * 0.02L = 3.132g.

E. Prepare 600ml of buffer A +0.5M NaCl. Take 600ml buffer A and add to it: 0.5mol/L * 58.44g/mol * 0.6L = 17.532g NaCl.

Part 4

A. Harvest cells and suspend in 120ml of ice cold buffer A with 2*u*g/ml each of leupeptin (80*u*l 10mg/ml stock), pepstatin (268*u*l 3mg/ml stock), aprotinin A (80*u*l of 10mg/ml stock), and DNAase A (80*u*l of 10mg/ml stock) . Lyse by two passes through a French Press at 1000psi. Clear lysate by centrifugation 16 000g for 50 minutes. Pellet is then washed in 5ml/tube buffer A then centrifuged again at 13 700rpm for 50 minutes. Take supernatant and load onto a 80ml Heparin-Sepharose HiTrap column (Pharmacia), equilibrated at 4°C in buffer A. Make sure that the column is washed with buffer A until the base line is stable. PP-1c is eluted by applying a linear gradient to buffer A up to 0.5M NaCl totaling 400ml. Fractions (5ml) are collected at a flow rate of 5ml/min.

B. Prepare 4L of buffer B stock solution. Buffer B is 50mM imidazole, 0.5mM EDTA, 0.5mM EGTA, 3mM $MnCl_2$, 3mM DTT, 0.5mM PMSF, 2mM benzamidine, and 10% glycerol at pH 7.2. 0.05mol/L * 68.08g/mol * 4L = 13.616g imidazole. 0.0005mol/L * 372.24g/mol * 4L = 0.7445g EDTA. 0.0005mol/L * 380.4g/mol * 4L = 0.7608g EGTA. 0.003mol/L * 154.24g/mol * 4L = 1.8509g DTT. Set the pH to 7.2. Then want 0.003mol/L * 197.9g/mol * 4L = 2.3748g, so use 2.3748g * mol/197.9g * L/0.2mol = 60ml $MnCl_2$ stock. Want 0.0005mol/L * 174.19g/mol * 4L = 0.3484g PMSF, so use 0.3484g * mol/174.19g * L/0.75mol = 2.7ml stock PMSF. Want 0.002mol/L * 156.6g/mol * 4L =1.2528g benzamidine, so use 1.2528g * mol/156.6g * L/1mol = 8ml stock benzamidine. Add 400 ml glycerol. From this 4L amount take 1L and make 50mM NaCl by adding 2.922g (0.05mol/L * 58.44g/mol * 1L = 2.922g NaCl). Also take 500ml and make 0.4M NaCl by adding 11.688g (0.4mol/L * 58.44g/mol X 0.5L).

Part 5

A. Do PNPP Assay of most of the fractions and solutions. Substrate is 30mM p-Nitrophenyl phosphate, MW 181.1g/mol, and want 1-2ml. Buffer is 50mM Tris, 1mM EDTA, 34mM $MgCl_2$, 0.5mM $MnCl_2$, pH 8.3, 0.2% mercaptoethanol fresh, and 1mg/ml BSA. Stop solution is 2M Na_2CO_3. Blank is 50*u*l buffer, control is 40*u*l buffer and 10*u*l protein phosphatase. Samples are 48*u*l buffer and 2*u*l fraction or solutions 40*u*l buffer and 10*u*l solution. 10 minutes later add 10*u*l 30mM PNPP. 10 minutes later add 4*u*l 2M Na_2CO_3 stop solution. Active fractions (usually with a reading higher than 0.5) are pooled. The volume is then measured and the sample is diluted by adding **2X** more buffer B. Make sure that the concentration of NaCl does not go under 100mM or else precipation will occur. Take 10*u*l from each active and side pooled fractions for electrophoresis.

B. Load protein sample onto a HR 10/10 Mono Q (Pharmacia) anion exchange column equilibrated in buffer B and washed with buffer B until baseline is stable. Elute protein by applying a gradient of 160ml buffer B with 0.4M NaCl. Rerun wash and load eluent at least two more times through this column to obtain a large quantity of pure protein. Multiple dialysis in a 500mM NaCl dialysis buffer is required to remove the Mn so that it does not interfere with the NMR spectra.

Peptides Preparation

$^{15}N/^{13}C$ labelled microcystin-LR was a gift from Bayer (Dr. Klaus Frobel, Dr. Hartwig Muller, and Dr. Thomas Henkel). The free double labelled microcystin-LR sample used for NMR was 2mM dissolved in a 10mM potassium phosphate, 50mM NaCl buffer with 80-90% H_2O/10-20% D_2O at pH 7. 2,2-dimethyl-2-silapentane-5-sulfonic acid (0.1mM) was added as a ^1H NMR chemical shift standard. The bound double labelled microcystin-LR (to PP-1c) was at approximately 0.1-0.2mM in approximately 1:1 ratio with PP-1c (any free microcystin-LR would confound the spectra). Instead of a 50mM NaCl concentration, a 500mM NaCl concentration was used in the sample containing bound microcystin-LR to increase the solubility of PP-1c. Free microcystin-LR spectra were re-recorded at 500mM NaCl and no discernible difference was observed relative to the spectra at 50mM NaCl.

Chemical Shift Analysis

The ^{15}N and ^{13}C NMR spectra were recorded on 500MHz or 600MHz Varian NMR spectrometers at 25°C. ^{15}N and ^{13}C NMR resonance assignments of ^{15}N- and ^{13}C-HSQC two dimensional NMR spectra for free microcystin-LR were made using the previous ^1H assignments (Chapter 2). Free PP-1c was confirmed by ^1H one-dimensional NMR.

Results and Discussion

Table 1 shows the ^{15}N/^{13}C/^{1}H chemical shifts for free microcystin-LR at 25°C. The ^{1}H chemical shifts (Chapter 2) were used in making the assignments of the ^{15}N and ^{13}C-HSQC. Separate ^{13}C spectra were aquired to cover the two distinct ^{13}C chemical shift of ranges 0-80ppm and 110-150ppm.

Table 1- ^{1}H, ^{15}N, and ^{13}C Chemical Shift Assignments (25°C) for Microcystin-LR (ppm)

Atom	Proton	^{15}Nitrogen
Arg 4:HN	8.58	125.31
Leu 2:HN	8.37	119.93
Ala 1:HN	8.21	118.02
Glu 6:HN	8.18	105.58
Adda 5:HN	7.94	131.09
Masp 3:HN	7.81	117.90

Atom	Proton	^{13}Carbon
Adda 5:HM	7.37	131.21
Adda 5:HK	7.29	132.30
Adda 5:HG	6.33	140.47
Mdha 7:HB1	5.96	118.74
Adda 5:HB	5.59	127.12
Mdha 7:HB2	5.58	118.74
Adda 5:HE	5.55	138.63
Adda 5:HA	4.49	58.15
Ala 1:HA	4.46	52.53
Masp 3:HA	4.44	59.25
Leu 2:HA	4.33	56.51
Arg 4:HA	4.29	54.33
Glu 6:HA	3.91	58.70
Adda 5:HH	3.48	9.48
Mdha 7:HM	3.40	40.73
Adda 5:H5	3.26	59.95
Masp 3:HB	3.22	44.17
Arg 4:HD	3.14	43.23
Adda 5:H1	3.03	46.83
Adda 5:HI1	2.92	39.48
Glu 6:HG1	2.81	34.33
Adda 5:HF	2.75	41.51
Adda 5:HI2	2.74	39.40
Glu 6:HG2	2.58	34.33
Glu 6:HB1	2.03	29.17
Leu 2:HB1	1.99	30.11
Arg 4:HB1	1.96	42.06
Glu 6:HB2	1.85	29.17
Adda 5:H3	1.70	14.87
Leu 2:HG	1.65	27.14
Arg 4:HB2	1.62	42.06
Leu 2:HB2	1.53	30.11

Atom	Proton	^{13}Carbon
Arg 4:HG	1.53	27.37
Masp 3:HG	1.38	18.78
Adda 5:H2	1.04	16.90
Adda 5:H4	1.01	18.54
Leu 2:HD1	0.90	25.18
Leu 2:HD2	0.86	22.61

Figure 1 displays the ^{15}N-HSQC spectrum for free microcystin-LR. The proton and ^{15}N chemical shift values used for crosspeak identification are in Table 1. The labels correspond to the assignments from Table 1.

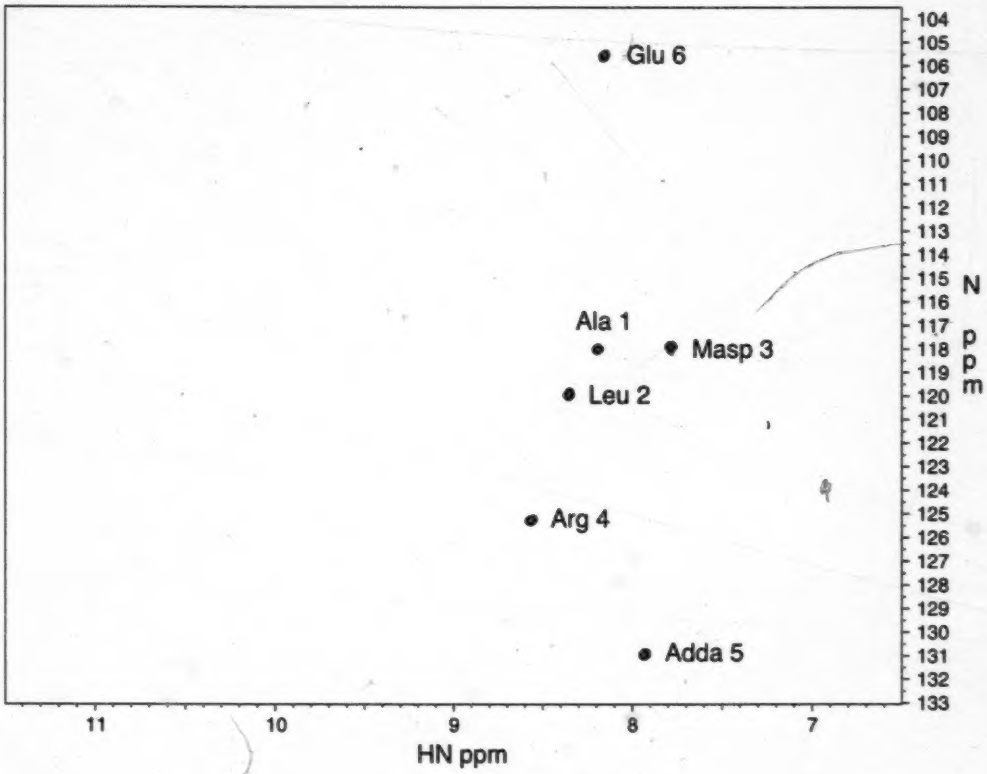


Figure 1-^{15}N-HSQC spectra for free microcystin-LR. Identification of peaks is from Table 1.

Figure 2 displays the same spectrum for microcystin-LR except now in the bound conformation. This ^{15}N-HSQC spectrum of bound microcystin-LR is of lesser quality relative the free form displayed in Figure 1. There are two reasons for the decrease in spectral quality. The first is the concentration of bound microcystin-LR is signficantly less than free concentration due to the precipitation of PP-1c. Secondly, broadening of lines (observed more in the ^{13}C-spectra e.g. Figure 4) is caused by an increase in molecular weight for microcystin-LR from 1 kDa (free) to 38 kDa (protein phosphatase-1 covalently linked to microcystin-LR). There are four discernible peaks in Figure 2 that are real. They have ^{15}N chemical shifts at approximately 113, 118, 120, and 130 ppms. The other peaks are noise or have an inverted phase.

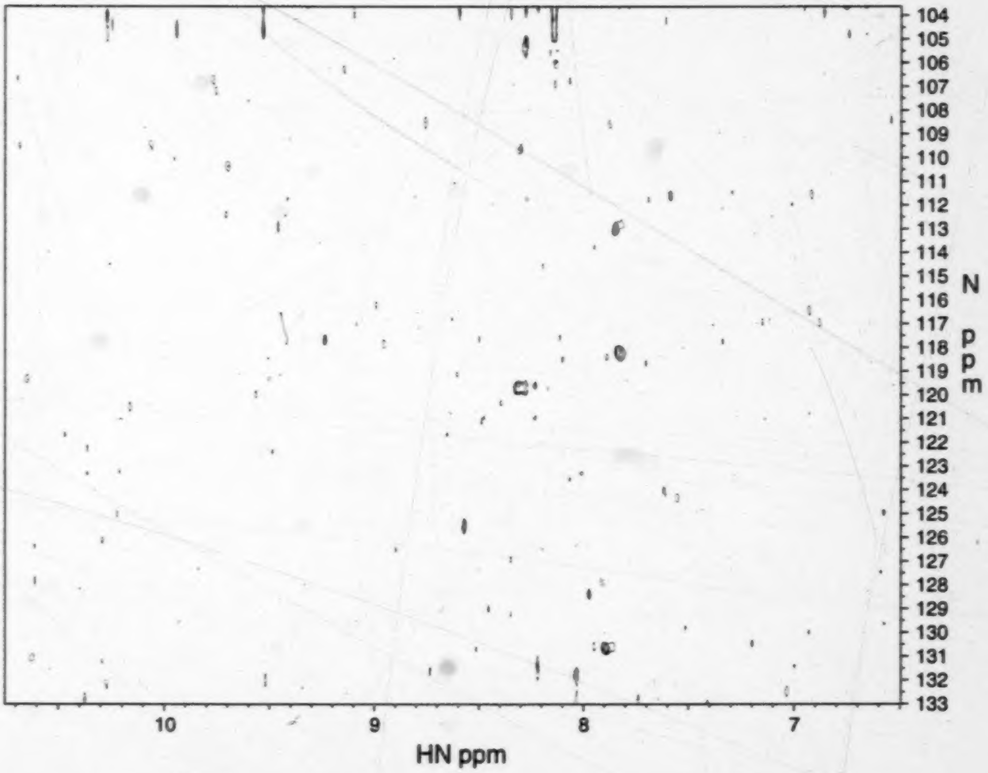


***Figure 2**-Bound ^{15}N-HSQC spectrum for microcystin-LR. This figure can be compared to Figure 1 for a direct comparison to free microcystin-LR.*

Figure 3 is the ^{13}C-HSQC spectrum (0-80ppm) for free microcystin-LR in the range of 0-80 ppm on the ^{13}C axis. Table 1 lists the identification of the crosspeaks with their specific chemical shifts.

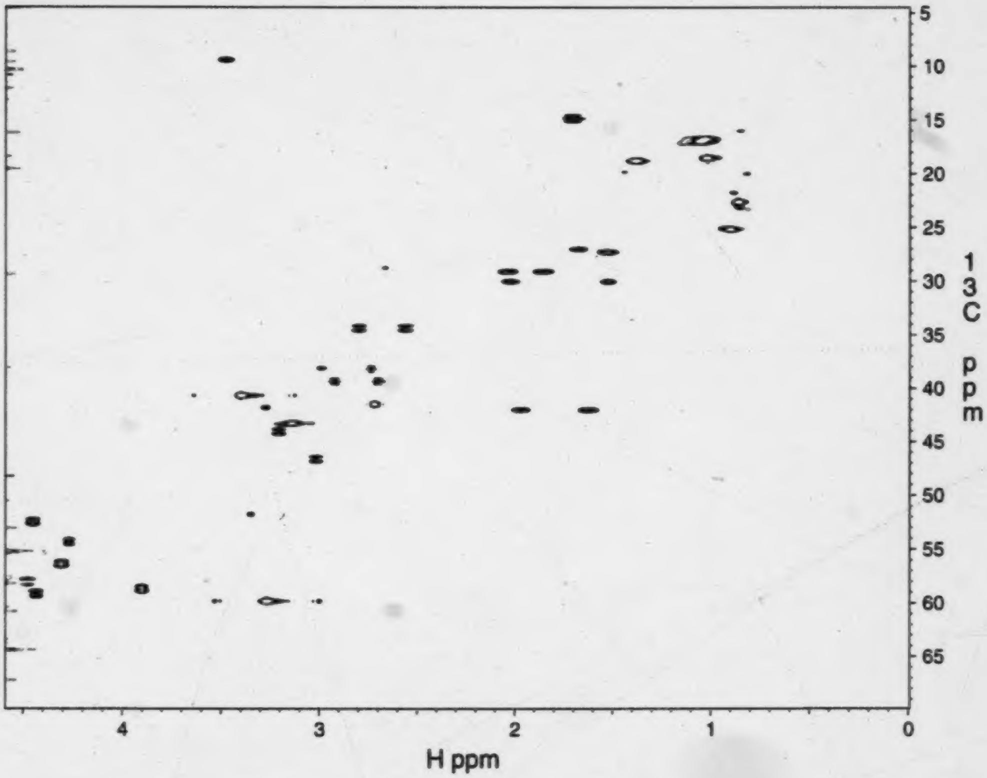


Figure 3-^{13}*C-HSQC spectrum for free microcystin-LR. Table 1 gives the chemical shifts and peak identification.*

Figure 4 presents the ^{13}C-HSQC (0-80ppm) spectrum for bound microcystin-LR. A direct comparison to the free form in Figure 3 shows how Figure 4 is missing peaks and has much broader crosspeaks.

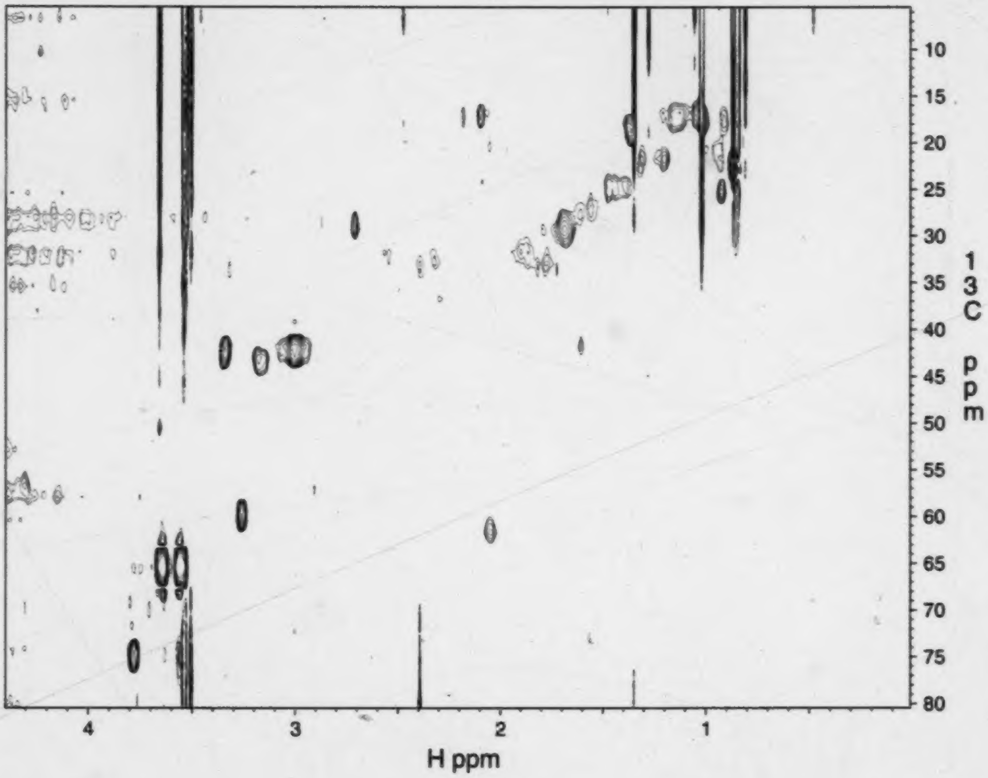


Figure 4-Bound microcystin-LR ^{13}C-HSQC spectrum. A direct comparison can be made to the free form in Figure 3.

Figure 5 is the ^{13}C-HSQC spectrum for free microcystin-LR (110-150ppm with respect to the ^{13}C axis). The specific values for the crosspeaks are given in Table 1 along with the identification of the crosspeak.

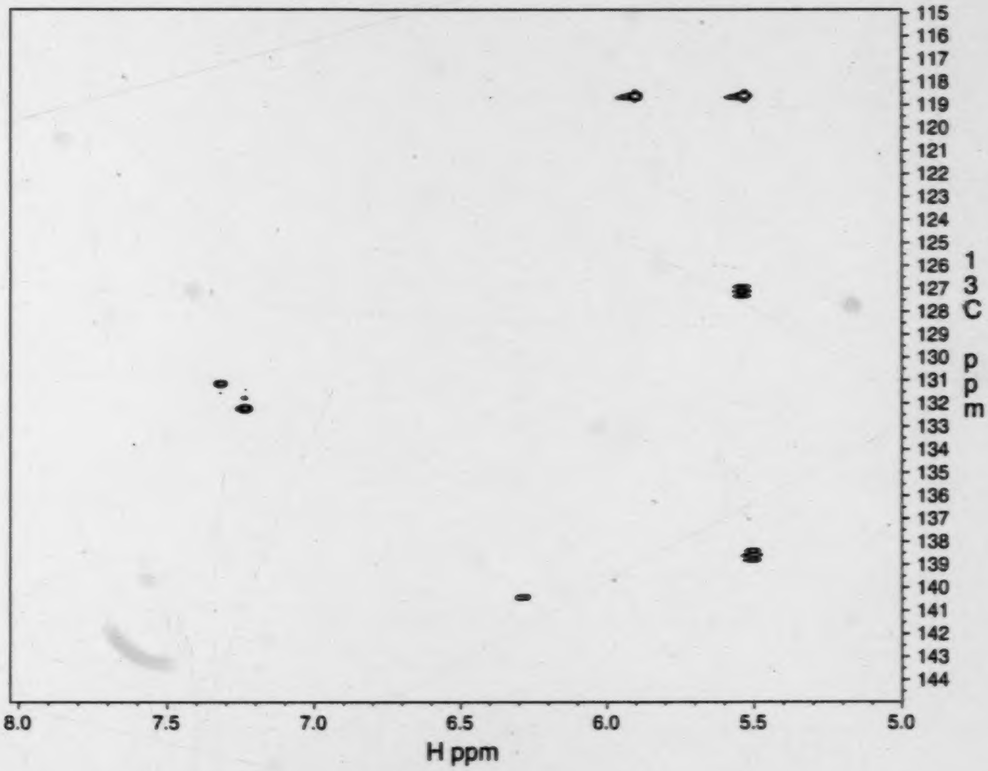


Figure 5- *Free microcystin-LR ^{13}C-HSQC spectrum above 80 ppm with respect to the ^{13}C axis. Table 1 gives the specific chemical shift values for the crosspeaks along with peak identification.*

Figure 6 is the ^{13}C-HSQC spectrum for bound microcystin-LR 110-150ppm with respect to the ^{13}C axis. There are no crosspeaks that are discernable. This again illustrates how the bound microcystin-LR spectra are unsuitable for structural determination in their present form.

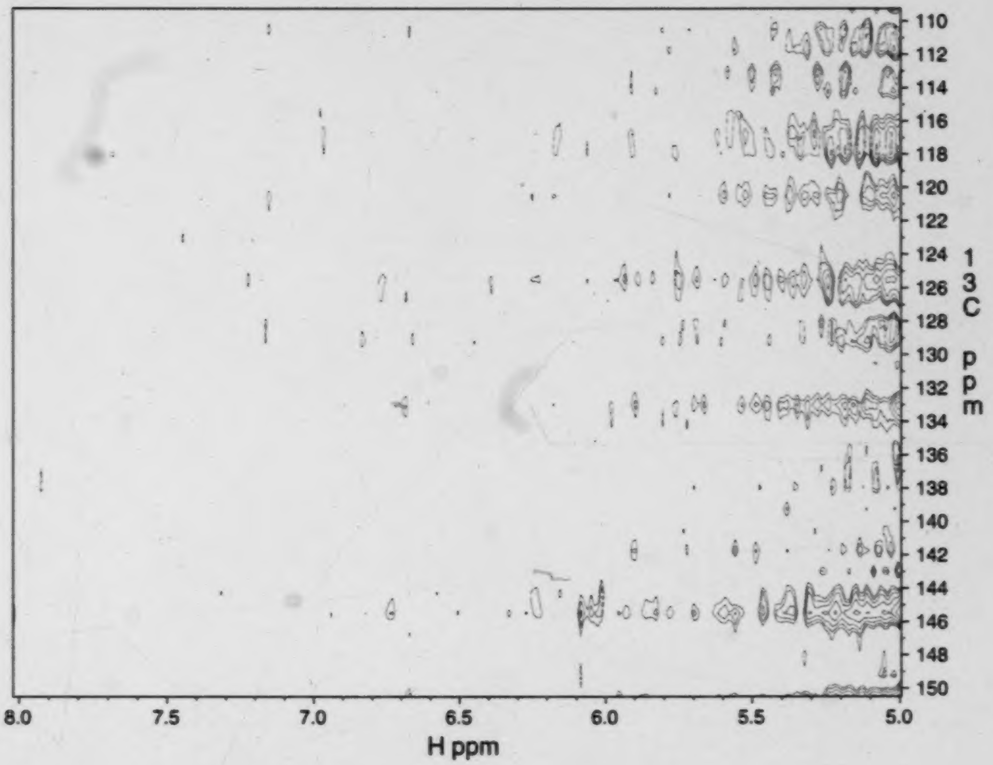


***Figure 6**-Bound microcystin-LR ^{13}C-HSQC spectrum above 80 ppm on the ^{13}C axis. A direct comparison to free microcystin-LR can be made using Figure 5.*

Preliminary binding experiments were unsuccessful because the ^{15}N- and ^{13}C-HSQC spectra turned out to be identical to the free microcystin-LR spectra due to the presence of a large excess of unbound microcystin-LR in the bound sample. This was solved by dialysizing out the free microcystin-LR leaving behind bound microcystin-LR:PP-1c. However, the bound ^{15}N and ^{13}C spectra turned out to be of very poor quality due to the solubility limitations of PP-1c. A variety of salts at various concentrations were used to try to increase the solubility of PP-1c to higher concentrations. The best salt was determined to be NaCl at a concentration of 500mM. The ^{15}N and ^{13}C spectra of free microcystin-LR at 50mM NaCl and 500mM NaCl had no significant differences. Even at 500mM NaCl PP-1c could only reach a concentration of 0.1-0.2mM. Unfortunately, at this concentration heavy precipitation occurred resulting in an even lower concentration of the bound microcystin-LR:PP-1c complex to a level where acceptable 15N and 13C two dimensional spectra could not be acquired.

Therefore, further investigation is required to increase the solutibility of PP-1c to a level where two-dimensional bound inhibitor NMR spectra can be acquired and analyzed. The experiments in this chapter were recorded using a 600 MHz spectrometer. It would be interesting to try a view a bound sample at a concentration below 0.1mM (current solubility level) on a 800 MHz spectrometer to determine if the increased sensitivity of the spectrometer causes acceptable two-dimensional spectra to be acquired.

References

1. Carmichael, W.W. The toxins of cyanobacteria. *Scientific American* **270(1)**, 78-86 (1994).

2. Honkanen, R.E. *et al.* Characterisation of microcystin-LR, a potent inhibitor of type-1 and -2A protein phosphatases. *J. Biol. Chem.* **265**, 19401-19404 (1990).

3. Yoshizawa, S. *et al.* Inhibition of protein phosphatases by microcystis and nodularin associated with hepatotoxicity. *J. Cancer Res. Clin. Oncol.* **116**, 609-614 (1990).

4. MacKintosh, C., Beattie, K.A., Klumpp, S., Cohen, P. & Codd, G.A. Cyanobacterial microcystin-LR is a potent and specific inhibitor of protein phosphatases 1 and 2A from both mammals and higher plants. *FEBS Lett.* **264**, 187-192 (1990).

5. Nishiwaki-Matsushima, R. *et al.* Structure-function relationships of microcystins, liver tumour promoters, in interaction with protein phosphatase. *Jpn. J. Cancer Res.* **82**, 993-996 (1991).

6. Holmes, C.F.B. & Boland, M.P. Inhibitors of protein phosphatase-1 and -2A; two of the major serine/threonine protein phosphatases involved in cellular regulation. *Current Opinion in Structural Biology* **3**, 934-943 (1994).

7. Eriksson, J.E. *et al.* Hepatocyte deformation induced by cyanobacterial toxins reflects inhibition of protein phosphatases. *Biochem. Biophys. Res. Commun.* **173**, 1347-1353 (1990).

8. Ohta, T. *et al.* Hyperphosphorylation of cytokeratins 8 and 18 by microcystin-LR, a new liver tumour promoter, in primary cultured rat hepatocytes. *Carcinogenesis* **13**, 2443-2447 (1992).

9. Bagu, J.R., Sönnichsen, F.D., Williams, D., Andersen, R.J., Sykes, B.D., and Holmes, C.F.B. *Nature Structural Biology* **2**, 114-116 (1995).

10. Goldberg, J., Huang, H., Kwon, Y., Greengard, P., Nairn, A.C., and Kuriyan, J. *Nature* **376**, 745-753 (1995).

Chapter 8

Conclusion

Part 1- Overview of Accomplishments

The main hypothesis of this thesis is that the marine toxins microcystin, nodularin, okadaic acid, and calyculin A and the inhibitory protein inhibitor-1 all have common structural features that allow these ligands to inhibit protein phosphatase-1. This hypothesis was tested by determining, modelling, or using previously published structures for both free and bound structures for all the inhibitors with the exception of inhibitor-1. In general, this hypothesis has been proven to be correct with the identification of common three-dimensional structural features for microcystin-LR, nodularin-V (motuporin), okadaic acid, and calyculin A and with the determination and/or modelling of these marine toxins at the same binding site on protein phosphatase-1. Unfortunately, inhibitor-1 turned out to have no distinct structure in the free state which prevented any comparisons to the other inhibitors or modelling onto protein phosphatase-1. However, useful structural data was acquired which resulted in the discovery that the inactive dephospho form of inhibitor-1 closely resembles the active phospho form. Finally, a large amount of data was gathered and analysized for microcystin-LR and microcystin-LL bound or modelled onto protein phosphatase-1 using structural thermodynamic calculations (STC).

The secondary hypothesis A that there is some structural difference between protein phosphatase-1 and protein phosphatase-2B (calcineurin) that would provide an explanation for why the marine toxins bind and inhibit protein phosphatase-1 while relatively being unable to inhibit calcineurin was also generally proven to be true. Although this structural difference caused by the connection between Asp 313:His 339 in calcineurin which displaces an important binding loop relative to its position in protein phosphatase-1 has yet to be tested with mutagenesis studies the complete analysis of all other structural differences which failed to have any significant affect upon toxin binding would suggest that this explanation is the only one available which accounts for the difference between the two enzymes.

There are many novel research accomplishments included in this thesis. These will now be covered on an individual basis:

1. The determination of the first free solution structure of microcystin-LR using nuclear magnetic resonance was published in Nature Structural Biology[1] (Chapter 2). This solution structure was later referred to by Goldberg et al.[2] in Nature when they published the crystal structure complex of microcystin-LR bound to protein phosphatase-1. A direct comparison between the average free solution structure[1] and the bound crystal

structure[2] of microcystin-LR which were found to be highly similar was published in the Journal of Biological Chemistry[3] (Chapter 3).

2. The determination of the first free solution structure of motuporin (nodularin-V) using nuclear magnetic resonance published in Nature Structural Biology along with microcystin-LR[1] (Chapter 2). A comparison[1] of this solution structure was made to the free structure of microcystin-LR in order to explain common binding elements. The significant difference in the three-dimensional positioning of the Mdha residue relative to the Mdhb residue provided a new explanation to why microcystin-LR is able to covalently link to protein phosphatase-1 while motuporin lacks this ability.

3. The determination of the first free solution structure of microcystin-LL using nuclear magnetic resonance published in the Journal of Biological Chemistry[3] (Chapter 3). Its similar structure to microcystin-LR was identified in this paper despite the change in residues from a hydrophilic arginine in microcystin-LR to a hydrophobic leucine in microcystin-LL.

4. Publication of the first models for individually bound okadaic acid and calyculin A to protein phosphatase-1 in the Journal of Biological Chemistry[3] using previously published crystal structures[2,4,5] (Chapter 3). These models were useful in showing how such diverse toxins in terms of primary structures could have similar three-dimensional structures that were able to inhibit protein phosphatase-1 by binding at a highly similar position. These models also confirmed previous biochemical data which examined the how inhibition was thought to occur.

5. Submission for publication to Protein Science of the structural thermodynamic calculation (STC) analysis for microcystin-LR bound to protein phosphatase-1 (Chapter 4). This represents the most important research in the thesis because it provides a novel approach for studying ligand:protein complexes. Although structural thermodynamic theory and procedures had been developed previously[6,7], the work presented in this thesis was novel because it applies to complexes that have not only been determined by crystallography but have also been modelled by taking a large number of free NMR solution structures and docking them onto a crystal structure. Since a large number of complexes were docked (e.g. approximately 50 or more) a more detailed understanding of ligand:protein binding was achieved in terms of microcystin-LR and -LL binding to protein phosphatase-1c. The basis for identifying poorly docked complexes due to their

low burial of accessible surface area and a higher dissociation constant was established. Most importantly, the entropy-enthalpy compensation principle where multiple complexes can have the similar free energies of binding despite having different entropy and enthalpies of binding was confirmed. Any differences in entropy when comparing two complexes of similar free energies of binding is offset by the corresponding change in the enthalpy and vice-versa. This was clearly observed when examining the multiple complexes for both microcystin-LR:protein phosphatase-1c and microcystin-LL:protein phosphatase-1c structures. This principle highlights the very nature of ligand:protein binding as being a dynamic process where minor fluctuations in structure and position can occur.

6. Submission for publication of a comparison of protein phosphatase-1[2] and calcineurin[8] crystal structures to the Journal of Biological Chemistry (Chapter 5). The focus of this paper was to be the first to explain why microcystin-LR and other toxins inhibit protein phosphatase-1 while having relatively no inhibition activity towards calcineurin. The toxins are non-inhibitors of calcineurin because a significantly important binding loop in protein phosphatase-1 is moved out of position in calcineurin due a connection between Asp 313 and His 339.

7. Submission for publication to the Journal of Biological Chemistry comparing the inactive dephospho form of inhibitor-1 with the active phospho form (Chapter 6). Although the lack of secondary and tertiary structure was very dissappointing useful insights into the differences between the two forms was achieved by the analysis of their chemical shifts. Due to the lack of any major changes in chemical shifts at a distance from Thr 35 where phosphorylation occurs it can be concluded that inhibitor-1 does not significantly change upon phosphorylation. Therefore, the addition of a strong negative charge at Thr 35 allows inhibitor-1 to bind to and inhibit protein phosphatase-1.

8. The first NMR solution studies of $^{15}N/^{13}C$ microcystin-LR bound to protein phosphatase-1 (Chapter 7). Although the crystal structure of this complex had previously been determined[2] no known studies of this complex in solution had been published. Determination of the 15N and 13C chemical shifts of free microcystin-LR was accomplished. A significant increase in the amount of recombinant protein phosphatase-1c was achieved (e.g. 10mg to 100mg) by using higher NaCl concentrations and rerunning wash and load eluent samples through the anion exchange column. Unfortunately, protein phosphatase-1 lacks solubility at concentrations higher than

0.1mM making analysis of the bound microcystin-LR two-dimensional spectra impractical.

Part 2- Future Research Directions

The largest obstacle in any future structural studies involving protein phosphatase-1 and its toxins or inhibitors in solution is the solubility limitation of the enzyme. It takes approximately 18.5 mg of protein phosphatase-1 in 500 ul of solution to achieve a 1 mM concentration. At high NaCl concentrations at 0.5 M, which was found to increase solubility during the expression and purification of protein phosphatase-1 relative to the lower salt concentrations previously used, only 2-3 mg of protein phosphatase-1 could be concentrated in a volume of 500 ul. Unfortunately, precipation of the enzyme occurred at this concentration indicating that a stable sample of protein phosphatase-1 was not achievable at approximately 0.1 mM. This is the only major stumbling block though. Double labelled inhibitors (e.g. microcystin-LR) are available and expression/purification of protein phosphatase-1 at amounts near 100 mg can be achieved within a month.

What are possible ways to overcome the solubility problem of protein phosphatase-1? Stronger field strength NMR magnets (e.g. 800 MHz) which are becoming available to use. Less concentrated samples are required to collect the same quality of data as the field strength of the spectrometers increases. Although many different types of salts were explored at a variety of concentrations to try to increase solubility of protein phosphatase-1, non-aqueous solvents were not test. Using solvents other than water may take away from the biological relevance of a structure but this factor might be offset if the limited solubility of protein phosphatase-1 was increased. Use of detergent or detergent-like substances that might be able to isolate individual protein phosphatase-1 molecules might be useful in preventing precipitation thereby increasing solubility. TROSY experiments might provide useful data of protein phosphatase-1 at the current conditions. As well, deuteration of protein phosphatase-1 could make it more practical for NMR experiments.

If the solubility barrier of protein phosphatase-1 can be overcome allowing the enzyme to be studied in detail by solution NMR then there are many interesting and important research projects that can be undertaken. The first of course is to determine the solution structure of protein phosphatase-1 by NMR. At 37 kDa the structural determination of protein phosphatase-1 would represent a major research achievement in

NMR. Better magnetics, multiple-dimensional NMR, and better peak identification techniques are making it possible to have sufficient resolution in spectra for structural determination. The majority of protein phosphatase-1 secondary structure is β-strand[2] which may also make for easier identification of crosspeaks.

A second possible research project after protein phosphatase-1 of prime importance is determining the bound form of inhibitor-1. The lack of significant structure in both the inactive dephospho and active phospho forms of inhibitor-1 was a big disappointment. As such no structural comparison to the marine toxin or modelling of an inhibitor-1 structure onto protein phosphatase-1 could be achieved. It is still unknown if inhibitor-1 does bind in the same site as microcystin-LR or other toxins. However, the lack of structure in both free forms of inhibitor-1 could actually make determination of the bound form much more interesting. Inhibitor-1 tightly binds to protein phosphatase-1 at low nanomolar concentrations. How can such an unstructured free inhibitor snap into a highly structured bound state necessary to form a tight binding complex? Determination of the bound state of inhibitor-1 would show how. Structural thermodynamic calculations of this inhibitor-1:protein phosphatase-1 complex could be used to measure the accessible surface area buried upon complex formation. It would be expected that large amounts of surface area would have to be buried in order to achieve the tight binding observed and offset the huge amount of conformation entropy required to stabilize the flexible and relatively non-structured free form into a bound conformation. Acquiring an $^{15}N/^{13}C$ labeled sample of inhibitor-1 would be a necessity.

A third possible project would be determination of the bound form of microcystin-LR in solution by NMR which was attempted in Chapter 7. This would be useful in order to test the structure thermodynamic calculation results in Chapter 4. An ensemble of free solution structures of microcystin-LR were individually docked onto protein phosphatase-1. The complexes were then analyzed by structural thermodynamic calculations. It would be interesting to calculated an ensemble of NMR structures of microcystin-LR in the bound form and see how they compare to docked structures both structurally and thermodynamically. Double labelled $^{15}N/^{13}C$ microcystin-LR is available.

Determining the bound conformation of motuporin (nodularin-V) would make for a fourth potential research project. The major point of interest in this case would be to confirm the hypothesis that the Mdhb residue of motuporin is nowhere near Cys-273 of protein phosphatase-1 which is why this toxin unlike microcystin-LR is unable to covalently link with the enzyme. An $^{15}N/^{13}C$ sample of motuporin would have to be made up but that should be possible to do without any major problems.

The fifth possible research project also involves marine toxins, namely okadaic acid and calyculin A. While their crystal structures have been determined[4,5] their structure in solution is unknown. Other than the models presented in Chapter 3 no bound information is available. Of the two toxins, okadaic acid might be the most difficult to work with using NMR because of it is a poly-ether fatty acid which may make solubility and spectral overlap a problem. Large amounts of okadaic acid and calyculin A in milligram quantities would have to be obtained.

A final research project would involve experimentally determining the heat capacity for the microcystin:protein phosphatase-1 complex using ITC. This would make the results from the STC chapter more meaningful if the calculated and experimental values were similar. It could also be applied to other protein phosphatase-1 inhibitors such as inhibitor-1.

The abundant number of important possible research projects indicates how the structural and functional examination of serine/threonine protein phosphatases, their inhibitors, and their toxins has still many important goals to be achieved. This field has progressed from a point at the beginning of this thesis where little relevant structural information was available to a position where important structures both free and bound have been determined. With their abundance and the wide variety of cellular systems that protein phosphatases regulate, it can be concluded that many major new discoveries will also be made in the future. Serine/threonine phosphatases role in human health especially as it relates to cancer will also be of major interest in the future. This clearly is a major research field in science that will be investigated for many years to come.

References

1. Bagu, J.R., Sönnichsen, F.D., Williams, D., Andersen, R.J., Sykes, B.D., and Holmes, C.F.B. (1995) *Nature Structural Biology* **2**, 114-116.

2. Goldberg, J., Huang, H., Kwon, Y., Greengard, P., Nairn, A.C., and Kuriyan, J. (1995) *Nature* **376**, 745-753.

3. Bagu, J.R., Sykes, B.D., Craig, M.M., and Holmes, C.F.B. (1997). *J. Biol. Chem.* **272(8)**, 5087-5097.

4. Tachibana, K., Scheuer, P.J., Tsukitani, Y., Kikuchi, H., Van Engen, D., Clardy, J., Gopichand, Y., Schmitz, F.J. (1981) *J Am Chem Soc* **103**, 2469-71.

5. Kato, Y., Fusetani, N., Matsunaga, S., Hashimoto, K., Fujita, S., and Furuya, T.J. (1986) *J Am Chem Soc* **108,** 2780-1.

6. Gomez J. and Freire E. (1996) *J Mol Biol 252*, 337-350.

7. Murphy K.P. and Freire E. (1992) *Adv Protein Chem 43,* 313-361.

8. Griffith, J.P., Kim, J.L., Kim, E.E., Sintchak, M.D., Thomson, J.A., Fitzgibbon, M.J., Fleming, M.A., Caron, P.R., Hsiao, K., and Navia, M.A. (1995) *Cell* **82,** 507-522.

Appendix 1-Using STC

Figure 1 is the main menu of STC. In order to perform structural thermodynamic calculations on a ligand:protein complex (e.g. attempting to dock MLR onto calcineurin is the example used in the appendix) one selects "binding" then presses "Calc ASA".

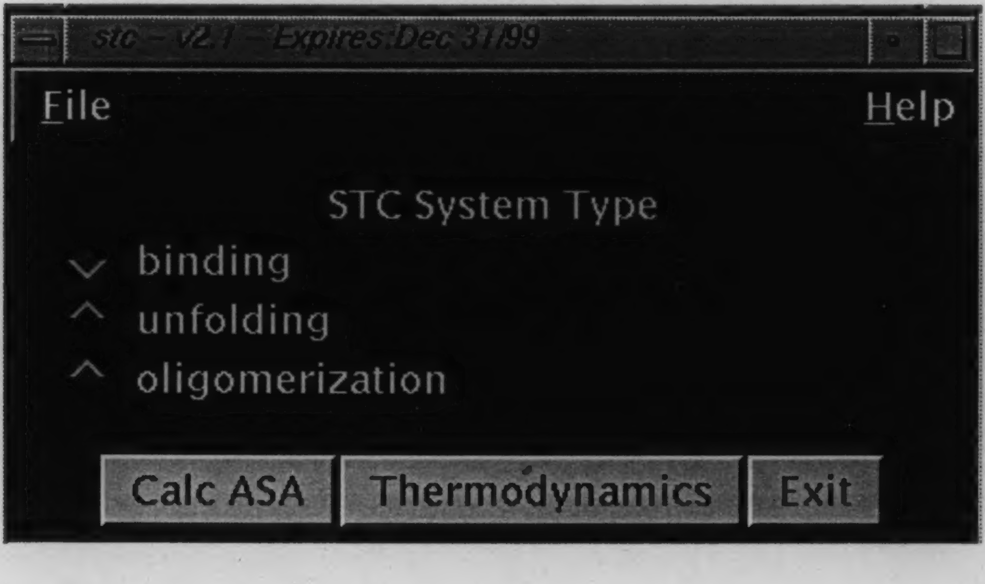
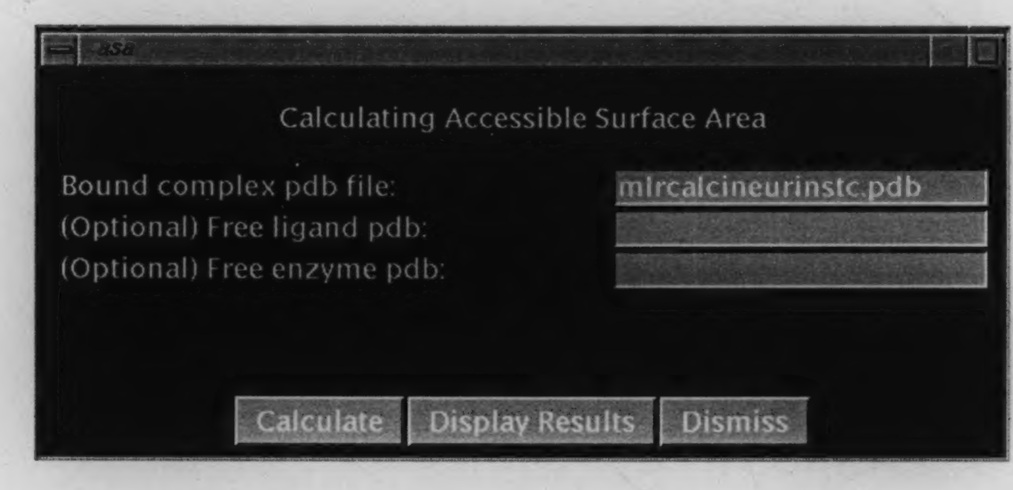


Figure 1- Main menu of STC.

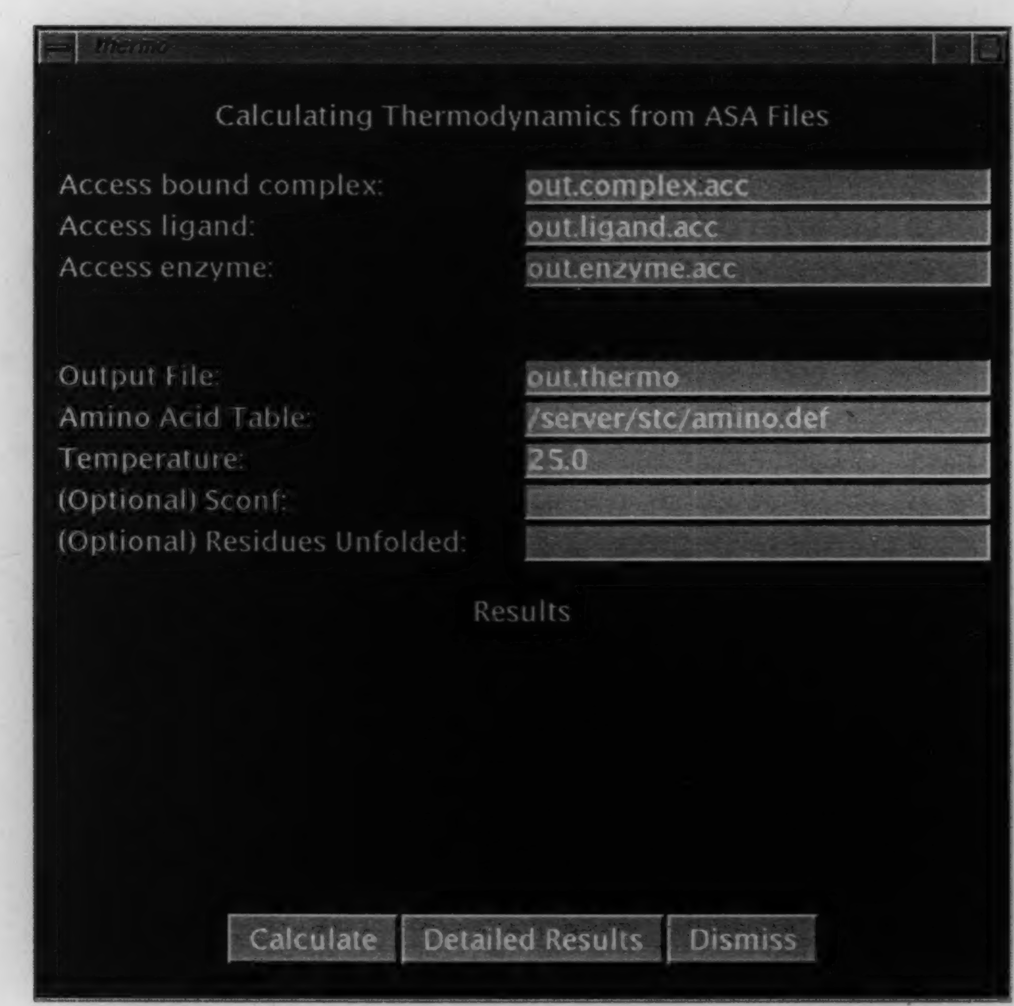
After pressing "Calc ASA" a new window appears (Figure 2). This window is where the pdb coordinates for the complex (free forms are optional) are loaded into the STC program. After the pdb file is typed in, "Calculate" is selected to continue the process. STC now calculates the nonpolar and polar ASA of the free and bound forms for the ligand and protein individually using the pdb coordinates. The free form ASA is subtracted from the bound form ASA to calculate the change in ASA for both the ligand and the protein.



Bound complex pdb file: mlrcalcineurinstc.pdb
(Optional) Free ligand pdb:
(Optional) Free enzyme pdb:

Figure 2- *Selecting pdb file to be used by STC by typing file name in.*

After STC calculates all the ASAs a person then goes back to the window containing the main menu (Figure 1) and selects "Thermodynamics". A new window appears (Figure 3) which displays the standard files used for calculating the thermodynamic parameters based on changes in nonpolar and polar ASA. To continue the process select "Calculate".



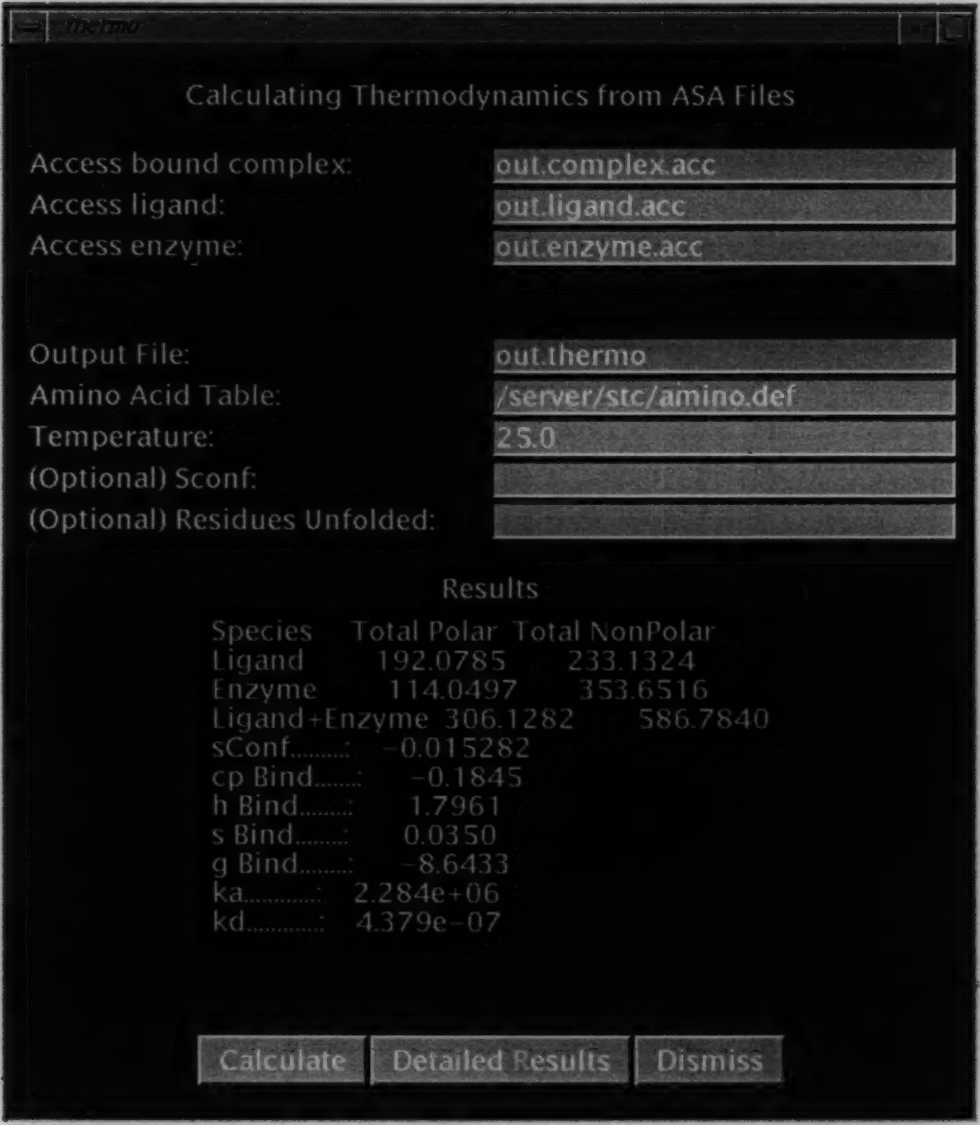
Calculating Thermodynamics from ASA Files

Access bound complex: out.complex.acc
Access ligand: out.ligand.acc
Access enzyme: out.enzyme.acc

Output File: out.thermo
Amino Acid Table: /server/stc/amino.def
Temperature: 25.0
(Optional) Sconf:
(Optional) Residues Unfolded:

Figure 3- Menu containing the standard files used in the thermodynamic calculations.

After pressing "Calculate" the general ASA and thermodynamic results are displayed within the window (Figure 4). To exit STC press "Dismiss" or "Exit" on all the menus. Detailed results are given in the default output file out.thermo.



Calculating Thermodynamics from ASA Files

Access bound complex: out.complex.acc
Access ligand: out.ligand.acc
Access enzyme: out.enzyme.acc

Output File: out.thermo
Amino Acid Table: /server/stc/amino.def
Temperature: 25.0
(Optional) Sconf:
(Optional) Residues Unfolded:

Results

```
Species      Total Polar  Total NonPolar
Ligand         192.0785      233.1324
Enzyme          114.0497      353.6516
Ligand+Enzyme  306.1282          586.7840
sConf.........:   -0.015282
cp Bind.......:     -0.1845
h Bind........:     1.7961
s Bind........:     0.0350
g Bind........:    -8.6433
ka............:  2.284e+06
kd............:  4.379e-07
```

Figure 4- Menu displaying the general ASA and thermodynamic results.